DISCRETE PROBABILITY DISTRIBUTIONS AND THEIR RECURSIVE RELATIONS

By

ROSE WANJIRU MUTHEE
I56/72731/2008

This project is submitted in partial fulfillment for the Degree of Master of Science in Mathematical Statistics in the School of Mathematics

UNIVERSITY OF NAIROBI

JULY, 2010

This project is my original work and has not been presented for a degree in any other University.

Signature...

R.W.MUTHEE

This project has been submitted for examination with my approval as the University Supervisor.

Signature...

PROF. J.A.M. OTTIENO

Acknowledgement

I desire to express my appreciations to those who have assisted me at different stages in the preparation of this dissertation.

First I thank the Almighty God for giving me good health throughout the period of study.

My thanks are due to my supervisor, Professor J.A.M Ottieno, for the constant encouragement, instructive discussions and valuable suggestions he has given me during the preparation of this Project and also Prof P.G.O.Weke and Dr. Nguti for their contribution.

To the members of staff School of Mathematics, University of Nairobi, I wish to express my gratitude for their support.

Also my thanks to my parents (Mr. Francis M. Mwangi and Mrs. Mercy W. Mwangi), my brother and sisters for their moral and financial support and their constant encouragement during my period of study.

Also my thanks goes to my fellow students; Tonui, Omukami, Otieno and Mecha for their supportive and interactive discussions throughout the period of study.

Finally I express my sincere thanks to my dear boyfriend Dann Mwaura for the constant moral support he has shown me throughout the period of study.

DEDICATION

I dedicate this project to my parents for the love and support they have shown me since I was born.

Abstract

The objective of this project is to study discrete probability distributions and their recursive patterns.

In Chapter I, we state the importance of expressing probability distributions in terms of recursive relations. This is because in some certain probability distributions it is often easier to deal with the recursive relations rather than the distributions themselves in obtaining the moments.

In Chapter II, we reviewed the various methods for determining discrete probability distributions. Some of these methods include both the binomial and exponential expansions, the Jacobian transformation (Change of Variable technique) .We also applied the expectation and convolution approaches to sums of iid random variables to obtain the resulting compound distributions.

In Chapter III, we have derived the recursive ratios and recursive relations for the various probability distributions that have been identified in Chapter II. Using the recursive relations, we have obtained the means and variances, based on the pgf technique (where possible) and Feller's method.

In Chapter IV, we reviewed a number of patterns of recursive relations; the main ones being the Panjer (1981) and Willmot (1988) patterns .With these patterns, we have been able to identify the corresponding probability distributions.

In Chapter V, we have reviewed some maximum likelihood estimation procedures. We have applied Sprott's procedure for deriving maximum likelihood equations.

Chapter VI contains the conclusions of this project and recommendations for further research.

Contents

Chapter 1

GENERAL INTRODUCTION

1.1 Problem Statement

One major study in statistics is Probability Distributions which can be discrete,continuous or mixed in nature.

In developng these distributions,some are not easily expressible can be expressed in easily expressible forms but their Probability Generating Functions and Recurrence Relations are.

Parameters of Probability Distributions and Recurrence Relations need to be estimated if not known.

Thus a number of issues arise in studying probability distributions.

1.2 Objectives

The main objectives of this project are:

- To review discrete probability mass functions
- To obtain recursive relations for these probability mass functions
- To study patterns of recursive relations
- To estimate parameters by Maximum Likelihood

Specific objectives are:

- To review Standard Discrete Distributions,Distributions based on Convolutions and Compound Distributions.

- To obtain the ratio $\frac{p_k}{p_{k-1}}$,where $p_k = Prob(X = k)$ hence to derive E(X) and Var(X) using pgf technique and Feller's method on the recursive relations obtained.

- To review patterns of recursive relations that have been developed;and identify probability mass functions corresponding to each pattern.

- To estimate the parameters of a convolution by Maximum Likelihood; to estimate the parameters of Compound distributions; to determine discrete distribution estimators from the recurrence equations for probabilities.

1.3 Application

The theories developed here can be applied in :

- Acturial Science for Insurance claims.

- Group-screening Designs for Identification and Estimation problems.

- Demography when studying waiting time birth-interval distributions in Human Reproductive Processes.

1.4 Literature Review

Let

$$X = Y_1 + Y_2 + \ldots + Y_N$$

denote the aggregate(total) claims amount where $X = 0$ if $N = 0$.
It is assumed that the severities Y_1,Y_2,... are mutually independent and distributed on the non-negative integers with common probability function

$$f_y = Prob(Y_i = y) \ ; y = 0, 1, 2, \ldots$$

It is further assumed that N is stochastically independent of Y_1,Y_2,... with probability function

$$p_n = Prob(N = n) \ ; n = 0, 1, 2, \ldots$$

If

$$g_x = Prob(X = x)$$

then,Feller (1968,Vol 1,p286-287) has shown that

$$g_x = \sum_{n=0}^{\infty} p_n f_n^{*n} \qquad (1.1)$$

where f^{*n} denotes the n-th convolution of f.

Using the pgf technique,

$$G(s) = P[F(s)] \qquad (1.2)$$

where F(s),P(s) and G(s) are pgfs of Y_i,N and X respectively.

Formula (1.1) may be difficult to use because of the high order of convolution.The computational complication of the convolutions prompted people to look for other alternative methods.

Panjer(1981) derived his famous recursive formula for the case where counting probabilities p_n satisfy the recursive relation

$$p_n = (a + \frac{b}{n})p_{n-1} \;; n = 1, 2, \ldots \qquad (1.3)$$

Under these assumptions,

$$g(x) = \sum_{i=1}^{x} (a + \frac{i}{x})p_i g(x - i)$$

Hess et al (2002) extended (1.3) to

$$p_{n+1} = (a + \frac{b}{n+1})p_n \;\; for \;\; n \geq k \qquad (1.4)$$

where

$$p_n = 0 \;\; for \;\; n \leq k - 1 \qquad (1.5)$$

which is called Panjer distribution of order k;with parameters a,b $\in \Re$ and k $\in \aleph$

Sundt and Jewel(1981) have shown a distribution that satisfies (1.3)(i.e. Panjer class of order 0) is

(i)Binomial if $a < 0$
(ii)Poisson if $a = 0$
(iii)Negative Binomial if $a > 0$

Willmot(1988)and Hess et al(2002) have identified all distributions of the Panjer class of order k with arbitrary k.

Recursions for a different extension of the class (1.3) can be found in Schroter(1990) and Sundt(1992).

Schroter(1990) extended Panjer's original recursion to

$$p_n = (a + \frac{b}{n})p_{n-1} + \frac{c}{n}p_{n-2} \; ; n = 1, 2, \ldots \tag{1.6}$$

with $p_{-1} = 0$
which is further generalized by Sundt(1992)

$$p_n = \sum_{j=1}^{k}(a_j + \frac{b_j}{n})p_{n-1} \; ; n = 1, 2, \ldots \tag{1.7}$$

with $p_n = 0$ for $n < 0$ In the same paper,Sundt(1992) extended (1.7) to

$$p_n = \sum_{j=1}^{k}(a_j + \frac{b_j}{n})p_{n-1} \; ; n = \omega + 1, \omega + 2, \ldots \tag{1.8}$$

However,the model fitting the class (1.8) in practical applications and the computational aspects are not discussed in Sundt(1992).Thus Panjer and Wang (1995) addressed these concerns.They stated that
(i)since it is desirable to try and fit a claim frequency model with relatively fair (2 or 3)parameters,the recursive relation (1.8) is useful only where k and ω are small.
(ii)a probability function can satisfy many different recursions.

In applications,among various recursive schemes,it would be good to know which one is preferable based on the following criteria
-stability
-simplicity
-computing effort.

There are many well known counting distributionswhich can fit into (1.8) with $k \leq 2$ and $\omega \geq 1$.

The Delaporte distribution (Ruohonen,1988;Willmot and Sundt,1989)which is in the class of Schroter(1990) satisfies (1.8) with $k = 2$ and $\omega = 0$.The

Polya-Aepli distribution (Johnson et al 1992,p329-330),which is not in the classes of Panjer(1981) or Schroter(1990),satisfies (1.8) with $k = 2$ and $\omega = 0$.

Other interesting examples for the general class of (1.8) can be found among the mixed Poisson distributions in Willmot(1993).

The Poisson-Pareto probability function satisfies (1.8) with $k = 2$ and $\omega = 2$.

The Poisson-Truncated normal satisfies (1.8) with $k = 2$ and $\omega = 0$.

Sundt(1992,p70-71) presents a nice argument on convolution of the members of the class of (1.7).He proves that the convolution of r distributions can be evaluated recursively as

$$p_n = \sum_{j=1}^{k}(a_j + \frac{b_j}{n})p_{n-j} \tag{1.9}$$

under certain conditions.

Panjer and Willmot(1982) went on to consider the class of counting distributions which satify a recursion

$$p_n = \frac{\sum_{t=0}^{k} a_t n^t}{\sum_{t=0}^{k} b_t n^t} p_{n-1} \quad ; n = 1, 2, \ldots \tag{1.10}$$

for some k, and derived recursions for the compound distribution when $k = 1$ and $k = 2$.These recursions were further developed by Willmot and Panjer(1987).

In the case of arbitrary k,it is not possible to give a complete characterization of the class (1.10).

Ord(1967) characterizes those distributions which satisfy a difference equation analogous to Pearson's differential equation,and also derives a recursive relation for the factorial moments.Also Guldberg(1931) considered recursive calculation of moments for certain members of the class (1.10).

Hesselager(1994) considers the class (1.10) and derives the following new recursive formula:

$$\begin{aligned} p_n \sum_{t=0}^{k} b_t n^t &= p_{n-1} \sum_{t=0}^{k} a_t n^t \\ &= p_{n-1} \sum_{t=0}^{k} c_t (n-1)^t \end{aligned}$$

Thus

$$p_n \sum_{t=0}^{k} b_t n^t = p_{n-1} \sum_{t=0}^{k} c_t (n-1)^t \tag{1.11}$$

where

$$c_t = \sum_{j=t}^{k} \binom{j}{t} a_j \tag{1.12}$$

Using (1.10),(1.11) and (1.12) Hesselager(1994) finds the parameters a_t,b_t and c_t;$t = 1, 2, \ldots, k$ for Waring (Beta-Geometric) distribution;the Polya-Eggenberger(Negative Hypergeometric) distribution;the hypergeometric distribution and the generalized Waring(Beta-Negative-Binomial) distribution.

Wang and Sobrero(1994) extended the recursive algorithm of Hesselager(1994) to a more general class of counting distributions,which includes Sundt's(1992) class as well as all the mixed Poisson distributions discussed by Willmot(1993).

The claim frequency N has a probability function satisfying

$$(\sum_{i=0}^{k} b_i n^i) p_n = \sum_{j=1}^{s} \{ \sum_{i=0}^{k} a_{j,i} (n-j)^i \} p_{n-j} \tag{1.13}$$

for $n = c, c+1, \ldots$ where c is a positive integer and $p_n = 0$ for $n < 0$.

Putting $k = s = c = 2$ in (1.13),Wang(1994) identified the matrix A

$$A = \begin{bmatrix} b_0 & b_1 & b_2 \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{bmatrix}$$

for

(i) the Sichel distribution which is obtained by mixing the Poisson mean over the Generalized Inverse Gaussian;

(ii) the Poisson-Beta distribution which is obtained by mixing the Poisson mean over the Beta distribution;

(iii) the Poisson Generalized Pareto which is obtained by mixing the Poisson mean over the Generalized Pareto;

(iv) the Poisson Inverse Gamma which is obtained by mixing the Poisson mean over the Inverse Gamma.

It is noted that Sundt's(1992) class is a particular case of (1.13) when $k = 1$.

1.5 Methodology

The methods used in this project are the pgf technique and Feller's method to obtain the moments (means and variances) of the distributions using the recursive relations.

For pgf techique we define

$$G(s) = \sum_{k=0}^{\infty} p_k s^k = \sum_{k=1}^{\infty} p_{k-1} s^{k-1}$$

$$G'(s) = \frac{dG}{ds} = \sum_{k=1}^{\infty} k p_k s^{k-1} = \sum_{k=2}^{\infty} (k-1) p_{k-1} s^{k-2}$$

$$G''(s) = \frac{d^2G}{ds^2} = \sum_{k=2}^{\infty} k(k-1) p_{k-1} s^{k-2} = \sum_{k=3}^{\infty} (k-1)(k-2) p_{k-1} s^{k-3}$$

$$E(X) = G'(1)$$

$$Var(X) = G''(1) + G'(1) - [G'(1)]^2$$

For Feller's method

$$M_1 = E(X) = \sum_{k=1}^{\infty} k p_k$$

$$M_2 = E(X^2) = \sum_{k=1}^{\infty} k^2 p_k$$

$$Var(X) = M_2 - M_1^2$$

Chapter 2

METHODS FOR DETERMINING DISCRETE PROBABILITY DISTRIBUTIONS

2.1 Introduction

Let X be a random variable taking the values 0,1,2,3,....
If $p_k = prob(X = k)$ implying that $0 \leq p_k \leq 1$
and $\sum_k p_k = 1$ then p_k is a probability mass function(pmf).

For a continuous random variable, a function f(x) is a probabilty density function (pdf) if
$f(x) > 0$ and $\int f(x)dx = 1$.

The aim of this chapter is to construct probability mass functions and probabilty density functions based on various methods.In particular,we shall deal with distributions based on exponential and binomial expansions,change of variable technique,sums of a fixed number of independent random variables and conditional probabilities given sums of two independent random variables.

We shall also construct compound distributions based on varying parameters of known probability distributions.Some of the parameters that become random variables will be of continuous form,namely,exponential,gamma and beta density functions.

Other compound distributions are based on the sum of a random number of independent random variables. It will be noticed that some compound distributions are easily expressed in terms of pgfs.However,their explicit probabilities are not easily obtained.

2.2 Distributions based on Exponential Expansion.

$$e^{\lambda} = 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^k}{k!} + \ldots \qquad (2.1)$$
$$= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

Therefore

$$1 = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}$$

Thus

$$p_k = \frac{e^{-\lambda}\lambda^k}{k!} \quad for \ k = 0, 1, 2, \ldots \qquad (2.2)$$

which is the Poisson probability distribution with parameter λ.

From equation (2.1)

$$e^{\lambda} - 1 = \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^k}{k!} + \ldots$$
$$= \sum_{k=1}^{\infty} \frac{\lambda^k}{k!}$$

Therefore

$$1 = \sum_{k=1}^{\infty} \frac{\lambda^k}{(e^{\lambda} - 1)k!}$$

Thus

$$p_k = (e^{\lambda} - 1)^{-1} \frac{\lambda^k}{k!} \ ; k = 1, 2, 3, \ldots \qquad (2.3)$$

which is called a zero-truncated Poisson distribution.

Equation (2.1) can also be expressed as

$$e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \sum_{k=0}^{j-1} \frac{\lambda^k}{k!} + \sum_{k=j}^{\infty} \frac{\lambda^k}{k!}$$

$$\Longrightarrow 1 = \sum_{k=0}^{j-1} e^{-\lambda}\frac{\lambda^k}{k!} + \sum_{k=j}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}$$

Therefore

$$1 - \sum_{k=0}^{j-1} \frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=j}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}$$

$$\Longrightarrow 1 = \frac{\sum_{k=j}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}}{1 - \sum_{k=0}^{j-1} \frac{e^{-\lambda}\lambda^k}{k!}}$$

Therefore

$$p_k = \frac{\frac{e^{-\lambda}\lambda^k}{k!}}{1 - \sum_{k=0}^{j-1} \frac{e^{-\lambda}\lambda^k}{k!}} \;; k = j, j+1, \ldots \tag{2.4}$$

and $p_k = 0$ for $k \leq j - 1$

which is called the Extended Poisson distribution.

2.3 Distributions based on Binomial Expansions.

Consider $(a + b)^n$.

When n is a positive integer,then

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} b^k a^{n-k} \tag{2.5}$$

Putting $b = p$ and $a = 1 - p$ then

$$1 = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

Therefore

$$p_k = \binom{n}{k} p^k (1-p)^{n-k} \;\; for \;\; k = 0, 1, 2, \ldots, n \tag{2.6}$$

which is a Binomial distribution with parameters n and p,denoted by b(n,p)

When n=1, and $b = p$ and $a = 1 - p$ then

$$p_k = p^k (1-p)^{1-k} \;\; for \;\; k = 0, 1 \tag{2.7}$$

which is a Bernoulli distribution with parameter p.

When n is a negative integer,let $n = -r$ where r is a positive integer.Then

$$(a+b)^{-r} = \sum_{k=0}^{\infty} \binom{-r}{k} b^k a^{-r-k}$$

Putting $a = 1$ and $b = -s$,we have

$$\begin{aligned} (1-s)^{-r} &= \sum_{k=0}^{\infty} (-1)^k \binom{-r}{k} s^k \\ &= \sum_{k=0}^{\infty} \binom{r+k-1}{k} s^k \end{aligned} \tag{2.8}$$

Putting $s = q$,we have

$$(1-q)^{-r} = \sum_{k=0}^{\infty} \binom{r+k-1}{k} q^k \tag{2.9}$$

Therefore

$$\begin{aligned} 1 &= \sum_{k=0}^{\infty} \binom{r+k-1}{k} q^k (1-q)^r \\ &= \sum_{k=0}^{\infty} \binom{r+k-1}{k} q^k p^r ; p = (1-q) \end{aligned} \tag{2.10}$$

Thus

$$p_k = \binom{r+k-1}{k} q^k p^r \;\; for \;\; k = 0, 1, 2, \ldots \tag{2.11}$$

which is a Negative Binomial Distribution where the random variable $X = k$ is the number of failures before the r^{th} success.

From (2.10)

$$1 = \binom{r-1}{0} q^0 p^r + \sum_{k=1}^{\infty} \binom{r+k-1}{k} q^k p^r$$
$$= p^r + \sum_{k=1}^{\infty} \binom{r+k-1}{k} q^k p^r$$

i.e.separating $k = 0$ from $k \geq 1$
Therefore we have

$$1 - p^r = \sum_{k=1}^{\infty} \binom{r+k-1}{k} q^k p^r$$
$$\Longrightarrow 1 = \sum_{k=1}^{\infty} \frac{p^r}{1-p^r} \binom{r+k-1}{k} q^k$$

Thus

$$p_k = \frac{p^r}{1-p^r} \binom{r+k-1}{k} q^k \ ; k = 1, 2, 3, \ldots \, and \ r = 1, 2, 3, \ldots \quad (2.12)$$

which is a zero-Truncated Negative Binomial Distribution.

Let X = the total number of trials required to achieve r successes
If

$$X = x = k + r$$

then

$$k = x - r$$

and (2.9) becomes

$$(1-q)^{-r} = \sum_{x-r=0}^{\infty} \binom{r+(x-r)-1}{x-r} q^{x-r}$$
$$= \sum_{x-r=0}^{\infty} \binom{x-1}{x-r} q^{x-r}$$
$$\Longrightarrow 1 = \sum_{x-r=0}^{\infty} \binom{x-1}{x-r} q^{x-r} p^r$$

Thus

$$p_k \;=\; \binom{x-1}{x-r} q^{x-r} p^r \;\; for \;\; x = r, r+1, r+2, \ldots \tag{2.13}$$

which is a Negative Binomial Distribution with the random variable $X = x$ being the total number of trials required to achieve r successes.

Again from (2.9)

$$(1-q)^{-r} = \sum_{k=0}^{\infty} \binom{r+k-1}{k} q^k = \sum_{k=0}^{\infty} (-1)^k \binom{-r}{k} (-q)^k$$
$$= \sum_{k=0}^{m-1} \binom{-r}{k} (-q)^k + \sum_{k=m}^{\infty} \binom{-r}{k} (-q)^k$$

$$\Longrightarrow (1-q)^{-r} - \sum_{k=0}^{m-1} \binom{-r}{k} (-q)^k = \sum_{k=m}^{\infty} \binom{-r}{k} (-q)^k$$

that is

$$(1-q)^{-r} - \sum_{k=0}^{m-1} \binom{r+k-1}{k} q^k = \sum_{k=m}^{\infty} \binom{r+k-1}{k} q^k$$

Therefore

$$1 = \frac{\sum_{k=m}^{\infty} \binom{r+k-1}{k} q^k}{(1-q)^{-r} - \sum_{k=0}^{m-1} \binom{r+k-1}{k} q^k}$$

Thus

$$p_k = \;=\; \frac{\binom{r+k-1}{k} q^k}{(1-q)^{-r} - \sum_{k=0}^{m-1} \binom{r+k-1}{k} q^k} \;\; ; x = m, m+1, \ldots \tag{2.14}$$

and $p_k = 0$ for $k \leq m-1$

which is called the Extended Negative Binomial Distribution.

When $s = q$ and $r = 1$ formula (2.8) becomes

$$(1-q)^{-1} \;=\; \sum_{k=0}^{\infty} q^k \tag{2.15}$$
$$1 \;=\; \sum_{k=0}^{\infty} q^k p$$

Therefore

$$p_k = q^k p \ for \ k = 0, 1, 2, \ldots \tag{2.16}$$

which is a Geometric distribution with parameter $p = 1 - q$ and where the random variable $X = k$ is the number of failures before a success.

Let $X = x = k + 1 \implies k = x - 1$

Then (2.15) becomes

$$(1 - q)^{-1} = \sum_{x-1=0}^{\infty} q^{x-1}$$

$$= \sum_{x=1}^{\infty} q^{x-1}$$

Therefore

$$1 = \sum_{x=1}^{\infty} q^{x-1} p$$

Thus

$$p_x = q^{x-1} p \ for \ k = 1, 2, 3, \ldots \tag{2.17}$$

which is a Geometric distribution with the random variable $X = x$ being the total number of trials required to achieve a success.

For truncated geometric distribution,we can re-arrange (2.15) as follows:

$$(1 - q)^{-1} = 1 + q + q^2 + \ldots + q^{m-1} + \sum_{k=m}^{\infty} q^k$$

$$= \frac{1 - q^m}{1 - q} + \sum_{k=m}^{\infty} q^k$$

$$\implies \frac{1}{1 - q} - \frac{1 - q^m}{1 - q} = \sum_{k=m}^{\infty} q^k$$

Therefore

$$\frac{q^m}{1-q} = \sum_{k=m}^{\infty} q^k$$
$$\Longrightarrow 1 = \sum_{k=m}^{\infty} q^k \frac{1-q}{q^m}$$
$$= \sum_{k=m}^{\infty} q^{k-m} p \;\; ; p = 1-q$$

Thus

$$p_k = q^{k-m} p \;\; for \;\; k = m, m+1, m+2, \ldots \qquad (2.18)$$

which is m-truncated Geometric distribution.

Equation (2.15) can also be re-arranged as

$$(1-q)^{-1} = \sum_{k=0}^{\infty} q^k = \sum_{k=0}^{m-1} q^k + \sum_{k=m}^{\infty} q^k \;\; for \;\; -1 < a < 1$$

Therefore

$$(1-q)^{-1} - \sum_{k=0}^{m-1} q^k = \sum_{k=m}^{\infty} q^k$$
$$\Longrightarrow 1 = \frac{\sum_{k=m}^{\infty} q^k}{(1-q)^{-1} - \sum_{k=0}^{m-1} q^k}$$

Thus

$$p_k = \frac{q^k}{(1-q)^{-1} - \sum_{k=0}^{m-1} q^k} \;\; ; k = m, m+1 \ldots \qquad (2.19)$$

and $p_k = 0$ for $k \leq m-1$
which is called the Extended Geometric distribution.

Again from (2.15)

$$\frac{1}{1-q} = 1 + q + q^2 + q^3 + \ldots$$
$$\Longrightarrow \int \frac{dq}{1-q} = \int [1 + q + q^2 + q^3 + \ldots] dq$$

$$\Longrightarrow -log(1-q) \;=\; q+\frac{q^2}{2}+\frac{q^3}{3}+\ldots=\sum_{k=1}^{\infty}\frac{q^k}{k} \tag{2.20}$$

Therefore

$$1=\sum_{k=1}^{\infty}\frac{q^k}{-klog(1-q)}$$

Thus

$$p_k \;=\; \frac{q^k}{-klog(1-q)} \;\; for \;\; k=1,2,3,\ldots \tag{2.21}$$

which is a Logarithmic distribution with parameter q.

The logarithmic series is given by

$$-log(1-q)=\sum_{j=1}^{\infty}\frac{q^j}{j}=\sum_{j=1}^{\infty}\frac{q^j}{\binom{j}{1}}$$

Extend $\frac{q^j}{\binom{j}{1}}$ to $\frac{q^j}{\binom{j}{m}}$ for $j \geq m$

Thus

$$\sum_{j=m}^{\infty}\frac{q^j}{\binom{j}{m}}=\sum_{k=m}^{\infty}\frac{q^k}{\binom{k}{m}}$$

i.e.

$$\sum_{j=m}^{\infty}\binom{j}{m}^{-1}q^j=\sum_{k=m}^{\infty}\binom{k}{m}^{-1}q^k$$

$$\Longrightarrow 1=\sum_{k=m}^{\infty}\left[\frac{\binom{k}{m}^{-1}q^k}{\sum_{j=m}^{\infty}\binom{j}{m}^{-1}q^j}\right]$$

Therefore

$$p_k \;=\; \frac{\binom{k}{m}^{-1}q^k}{\sum_{j=m}^{\infty}\binom{j}{m}^{-1}q^j} \;\; for \;\; k=m,m+1,\ldots \tag{2.22}$$

which is the Extended Logarithmic distribution.

2.4 Distributions based on Change of Variable Technique

2.4.1 Exponential Distribution

Let X be a continuous uniform distribution over the interval [0,1].Thus

$$f(x) = 1 \ , 0 \leq x \leq 1$$

and zero elsewhere.
If

$$Y = -lnX^{\frac{1}{\lambda}} \ , \lambda > 0$$

then

$$\begin{aligned} lnx^{\frac{1}{\lambda}} &= -y \\ \Longrightarrow x^{\frac{1}{\lambda}} &= e^{-y} \\ \Longrightarrow x &= e^{-\lambda y} \\ \Longrightarrow dx &= -\lambda e^{-\lambda y} dy \end{aligned}$$

Therefore the pdf of y is

$$\begin{aligned} g(y) &= f(x) \mid J \mid \\ &= 1. \mid \frac{dx}{dy} \mid \\ &= \mid -\lambda e^{-\lambda y} \mid \end{aligned}$$

$$g(y) = \lambda e^{-\lambda y} \qquad (2.23)$$

for $y > 0$
which is an exponential distribution with parameter λ.

2.4.2 Gamma Density Function

A gamma function of $\alpha > 0$ is defined by

$$\begin{aligned} \Gamma\alpha &= \int_0^\infty x^{\alpha-1} e^{-x} dx \\ \Longrightarrow 1 &= \int_0^\infty \frac{x^{\alpha-1} e^{-x}}{\Gamma\alpha} dx \end{aligned}$$

Therefore

$$f(x) = \frac{x^{\alpha-1}e^{-x}}{\Gamma\alpha} \tag{2.24}$$

for $x > 0$
is a gamma density function with one parameter α.

Let

$$y = \frac{x}{\beta}$$

$$\Longrightarrow x = \beta y$$

and

$$dx = \beta dy$$

The pdf of Y is given by

$$\begin{aligned} g(y) &= f(x) \mid J \mid \\ &= \frac{x^{\alpha-1}e^{-x}}{\Gamma\alpha} \mid \frac{dx}{dy} \mid \\ &= \frac{1}{\Gamma\alpha} e^{-\beta y} (\beta y)^{\alpha-1} \beta \end{aligned}$$

Thus

$$g(y) = \frac{\beta^{\alpha}}{\Gamma\alpha} y^{\alpha-1} e^{-\beta y} \tag{2.25}$$

for $y > 0$
is a gamma density function with two parameters α and β.

Now let

$$z = \beta x$$

$$\Longrightarrow x = \frac{z}{\beta}$$

and

$$dx = \frac{dz}{\beta}$$

Therefore the pdf of Z is

$$\begin{aligned} h(z) &= f(x) \mid J \mid \\ &= \frac{x^{\alpha-1}e^{-x}}{\Gamma\alpha} \mid \frac{dx}{dz} \mid \\ &= \frac{e^{-\frac{z}{\beta}}(\frac{z}{\beta})^{\alpha-1}}{\Gamma\alpha}\frac{1}{\beta} \end{aligned}$$

Thus

$$h(z) = \frac{e^{-\frac{z}{\beta}}z^{\alpha-1}}{\beta^{\alpha}\Gamma\alpha} \qquad (2.26)$$

for $z > 0$
This is another gamma density function with two parameters α and β.

2.4.3 Beta Density Function

Definition

A beta function of α and β is defined by

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \qquad (2.27)$$

for $0 < x < 1$
When we put

$$y = 1 - x$$

then
$x = 1 - y$ and $dx = -dy$
Therefore

$$\begin{aligned} B(\alpha,\beta) &= \int_0^1 (1-y)^{\alpha-1}y^{\beta-1}(-dy) \\ &= \int_0^1 y^{\beta-1}(1-y)^{\alpha-1}dy \end{aligned}$$

$$= B(\beta,\alpha) \qquad (2.28)$$

Next,put

$$x = \sin^2\theta$$
$$\Longrightarrow dx = 2\sin\theta\cos\theta d\theta$$

Therefore

$$B(\alpha,\beta) = \int_0^{\frac{\pi}{2}} (\sin\theta)^{2\alpha-2}(\cos\theta)^{2\beta-2} 2\sin\theta\cos\theta d\theta$$

$$= 2\int_0^{\frac{\pi}{2}} (\sin\theta)^{2\alpha-1}(\cos\theta)^{2\beta-1} d\theta \qquad (2.29)$$

Relationship between Beta and Gamma functions.

$$\Gamma\alpha\Gamma\beta = [\int_0^{\infty} x^{\alpha-1}e^{-x}dx][\int_0^{\infty} y^{\beta-1}e^{-y}dy]$$
$$= \int_0^{\infty}\int_0^{\infty} x^{\alpha-1}y^{\beta-1}e^{-(x+y)}dydx$$

Let $x = s^2$ and $y = t^2 \Longrightarrow dx = 2sds$ and $dy = 2tdt$
Therefore

$$\Gamma\alpha\Gamma\beta = \int_0^{\infty}\int_0^{\infty} s^{2\alpha-2}t2\beta - 2e^{-(s^2+t^2)}2sds2tdt$$
$$= 4\int_0^{\infty}\int_0^{\infty} s^{2\alpha-1}t2\beta - 1e^{-(s^2+t^2)}dsdt$$

Next let $s = r\sin\theta$ and $t = r\cos\theta$.

$$J = \begin{vmatrix} \frac{ds}{d\theta} & \frac{dt}{d\theta} \\ \frac{dt}{dr} & \frac{ds}{dr} \end{vmatrix}$$

i.e. J is the determinant of the matrix

$$A = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$= rcos^2\theta + rsin^2\theta = r$$

Thus

$$J = r$$

Therefore

$$\begin{aligned}
\Gamma\alpha\Gamma\beta &= 4\int_0^{\frac{\pi}{2}}\int_0^{\infty}(r sin\theta)^{2\alpha-1}(r cos\theta)^{2\beta-1}e^{-r^2} \mid J \mid dr d\theta \\
&= 2\int_0^{\infty} e^{-r^2}[2\int_0^{\frac{\pi}{2}}(sin\theta)^{2\alpha-1}(cos\theta)^{2\beta-1}d\theta]r^{2\alpha+2\beta-2}rdr \\
&= \int_0^{\infty} 2e^{-r^2}B(\alpha,\beta)r^{2(\alpha+\beta-1)}rdr
\end{aligned}$$

Put $u = r^2 \implies du = 2rdr$
Therefore

$$\begin{aligned}
\Gamma\alpha\Gamma\beta &= B(\alpha,\beta)\int_0^{\infty} r^{2(\alpha+\beta-1)}e^{-r^2}2rdr \\
&= B(\alpha,\beta)\int_0^{\infty} u^{\alpha+\beta-1}e^{-u}du \\
&= B(\alpha,\beta)\Gamma(\alpha+\beta)
\end{aligned}$$

Thus

$$B(\alpha,\beta) = \frac{\Gamma\alpha\Gamma\beta}{\Gamma(\alpha+\beta)} \tag{2.30}$$

But

$$\begin{aligned}
B(\alpha,\beta) &= \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \\
\implies 1 &= \int_0^1 \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}dx \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta}x^{\alpha-1}(1-x)^{\beta-1}dx
\end{aligned}$$

Therefore

$$f(x) = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta}x^{\alpha-1}(1-x)^{\beta-1} \tag{2.31}$$

for $0 < x < 1$
is a Beta pdf with parameters α and β.

2.5 Probability distributions based on sums of fixed number of iid random variables

Let

$$S_N = X_1 + X_2 + \ldots + X_N$$

where the $X_i's$ are iid random variables and N is fixed.
We wish to find the pgf of $S_N$using the Expectation Approach and the Convolution Approach and then apply it to the sum of standard discrete distributions.

2.5.1 Expectation Approach

Let

$$G_i(s) = E(s^{X_i})$$

the probability generating function (pgf) of X_i and

$$H(s) = E(s^{S_N})$$

the probability generating function (pgf) of S_N

$$\begin{aligned}
\Longrightarrow H(s) &= E(s^{S_N}) \\
&= E(s^{X_1+X_2+\ldots+X_N}) \\
&= E(s^{X_1}).E(s^{X_2})\ldots E(s^{X_N}) \\
&= G_1(s).G_2(s)\ldots G_N(s) \\
&= \prod_{i=1}^{N} G_i(s)
\end{aligned}$$

Since the $X_i's$ are identically distributed,then $G_i(s) = G(s)\ \forall i = 1, 2, \ldots, N$. Thus

$$\begin{aligned}
H(s) &= \prod_{i=1}^{N} G(s) \\
&= [G(s)]^N
\end{aligned}$$

2.5.2 Convolution Approach

Consider three sequences: $\{a_k\}$, $\{b_k\}$ and $\{c_k\}$.
If

$$
\begin{aligned}
c_k &= a_0 b_k + a_1 b_{k-1} + a_2 b_{k-2} + \ldots + a_r b_{k-r} + \ldots + a_k b_0 \\
&= \sum_{r=0}^{k} a_r b_{k-r}
\end{aligned}
$$

then we say $\{c_k\}$ is a convolution of $\{a_k\}$ and $\{b_k\}$ i.e.

$$\{c_k\} = \{a_k\} * \{b_k\}$$

Theorem

If the generating functions of $\{a_k\}$, $\{b_k\}$ and $\{c_k\}$ are:

$$A(s) = \sum_{k=0}^{\infty} a_k s^k$$

$$B(s) = \sum_{k=0}^{\infty} b_k s^k$$

and

$$C(s) = \sum_{k=0}^{\infty} c_k s^k$$

respectively,then

$$C(s) = A(s)B(s)$$

proof:

$$
\begin{aligned}
A(s)B(s) &= (a_0 + a_1 s + a_2 s^2 + \ldots)(b_0 + b_1 s + b_2 s^2 + \ldots) \\
&= a_0 b_0 + (a_0 b_1 + a_1 b_0)s + (a_o b_2 + a_1 b_1 + a_2 b_0)s^2 + \ldots \\
&= c_0 + c_1 s + c_2 s^2 + \ldots \\
&= \sum_{k=0}^{\infty} c_k s^k \\
&= C(s)
\end{aligned}
$$

Let X and Y be two independent random variables. If

$$p_i = Prob(X = i)$$

and

$$q_j = Prob(Y = j)$$

then

$$Prob(X = i, Y = j) = Prob(X = i)Prob(Y = j) = p_i q_j$$

Now let

$$Z = X + Y$$

Then

$$\begin{aligned} r_k &= Prob(Z = k) \\ &= Prob(X + Y = k) \\ &= Prob(X = 0, Y = k) + Prob(X = 1, Y = k - 1) + \ldots + Prob(X = k, Y = 0) \\ &= Prob(X = 0)Prob(Y = k) + Prob(X = 1)Prob(Y = k - 1) + \ldots + Prob(X = k)Prol \\ &= p_0 q_k + p_1 q_{k-1} + p_2 q_{k-2} + \ldots + p_k q_0 \end{aligned}$$

which implies that

$$\{r_k\} = \{p_k\} * \{q_k\}$$

i.e. $\{r_k\}$ is a convolution of $\{p_k\}$ and $\{q_k\}$

If $G_x(s)$,$G_y(s)$ and $G_z(s)$ are pgfs of $\{p_k\}$,$\{q_k\}$ and $\{r_k\}$ respectively,then from the theorem

$$G_z(s) = G_x(s)G_y(s)$$

Thus,the pgf of the sum of two independently distributed random variables is the product of two pgfs.

If

$$\{q_k\} = \{p_k\}$$

then

$$\{r_k\} = \{p_k\} * \{p_k\}$$

that is $\{r_k\}$ is a two-fold convolution of $\{p_k\}$with itself denoted by

$$\{r_k\} = \{p_k\}^{2*}$$

If X and Y are independently and identically distributed with common pgf G(s),then

$$G_z(s) = [G(s)]^2$$

Generalizing to an arbitrary number of random variables,we have the following theorem

Theorem 2.1

Let X_1,X_2,...,X_N be independent integer-valued random variables with probability distributions $\{p_{k,1}\}$,$\{p_{k,2}\}$,...,$\{p_{k,N}\}$ and pgfs $G_1(s)$,$G_2(s)$,...,$G_N(s)$ respectively. The probability distribution of the sum

$$Z_N = X_1 + X_2 + \ldots + X_N$$

which we denote by $\{r_k\}$ is the convolution of $\{p_{k,1}\}$,$\{p_{k,2}\}$,...,$\{p_{k,N}\}$ i.e.

$$\{r_k\} = \{p_{k,1}\} * \{p_{k,2}\} * \ldots * \{p_{k,N}\}$$

The pgf of Z_N is the product of $G_1(s)$,$G_2(s)$,...,$G_N(s)$;i.e.

$$G_{Z_N}(s) = G_1(s)G_2(s)\ldots G_N(s)$$

Corollary 2.1

If X_1,X_2,...,X_N are independent and identically distributed (iid) random variables with a common probability distribution $\{p_k\}$ and the same pgf G(s),then the probability distribution of the sum

$$Z_N = X_1 + X_2 + \ldots + X_N$$

is the n-fold convolution of $\{p_k\}$ with itself ; i.e.

$$\{r_k\} = \{p_k\}^{n*}$$

and the pgf is

$$G_{Z_N}(s) = [G(s)]^N$$

2.5.3 Application

2.5.3.1 The sum of Bernoulli variables

Given

$$p_k = Prob(X = k) = p^k(1-p)^{1-k} \quad ; k = 0, 1$$

then

$$G(s) = \sum_{k=0}^{1} p_k s^k = q + ps \qquad ; q = 1 - p$$

Hence

$$H(s) = [G(s)]^N$$
$$= [q + ps]^N$$

which is the pgf of a Binomial distribution with parameters N and p.
Thus the distribution of the sum of Bernoulli independent and identically distributed random variables is Binomial.

2.5.3.2 The sum of Binomial variables

Given

$$p_k = Prob(X = k) = \binom{m}{k} p^k (1-p)^{m-k} \quad ; k = 0, 1, 2, \ldots, m$$

then

$$G(s) = \sum_{k=0}^{m} p_k s^k = \sum_{k=0}^{n} \binom{m}{k} (ps)^k q^{m-k} \quad q = 1 - p$$
$$= (q + ps)^m$$

Hence

$$H(s) = [G(s)]^N$$
$$= [(q + ps)^m]^N$$
$$= (q + ps)^{mN}$$

which is the pgf of a binomial distribution with parameters Nm and p.
Thus the distribution of the sum of Binomial independent and identically distributed random variables is Binomial.

2.5.3.3 The sum of Geometric variables

Type I:

Let X be the number of failures before the first success.
Then

$$p_k = Prob(X = k) = q^k p \; ; k = 0, 1, 2, \ldots \quad ; q = 1 - p$$

and

$$G(s) = \sum_{k=0}^{\infty} p_k s^k = \sum_{k=0}^{\infty} (qs)^k p$$

$$= p \sum_{k=0}^{\infty} (qs)^k$$

$$= \frac{p}{1 - qs}$$

Hence

$$H(s) = [G(s)]^N$$

$$= [\frac{p}{1 - qs}]^N$$

which is the pgf of a Negative Binomial distribution with parameters N and p.

Type II:

Let X be the number of total number of trials required to achieve the first success.

Then

$$p_k = Prob(X = k) = q^{k-1}p; k = 1, 2, 3, \ldots \quad ; q = 1 - p$$

and

$$G(s) = \sum_{k=1}^{\infty} p_k s^k = \sum_{k=1}^{\infty} q^{k-1} s^k p$$

$$= ps \sum_{k=1}^{\infty} (qs)^{k-1}$$

$$= \frac{ps}{1 - qs}$$

Hence

$$H(s) = [G(s)]^N$$

$$= [\frac{ps}{1 - qs}]^N$$

which is also the pgf of a Negative Binomial distribution with parameters N and p.

Thus the distribution of the sum of Geometric independent and identically distributed random variables is Negative Binomial.

2.5.3.4 The sum of Negative Binomial variables

Type I:

Let X be the number of failures before the r^{th} success.
Then

$$p_k = Prob(X = k) = (-1)^k \binom{-r}{k} (1-p)^k p^r \; ; k = 0, 1, 2, \ldots$$
$$= \binom{r+k-1}{k} q^k p^r \quad ; q = 1 - p$$

and

$$G(s) = \sum_{k=0}^{\infty} p_k s^k = \sum_{k=0}^{\infty} (-1)^k \binom{-r}{k} (qs)^k p^r$$
$$= p^r \sum_{k=0}^{\infty} \binom{-r}{k} (-qs)^k$$
$$= \frac{p^r}{(1-qs)^r}$$

Hence

$$H(s) = [G(s)]^N$$
$$= [\frac{p}{1-qs}]^{Nr}$$

which is the pgf of a Negative Binomial distribution with parameters Nr and p.

Type II:

Let X be the number of total number of trials required to achieve the r^{th} success.
Then

$$p_k = Prob(x_i = k) = \binom{k-1}{r-1} (1-p)^{k-r} p^r \; ; k = r, r+1, r+2, \ldots$$

and

$$G(s) = \sum_{k=r}^{\infty} p_k s^k = \sum_{k=r}^{\infty} \binom{k-1}{r-1} q^{k-r} p^r s^k \quad ; q = 1 - p$$

$$= (ps)^r \sum_{k=r}^{\infty} \binom{k-1}{r-1} (qs)^{k-r}$$

$$= (ps)^r [\binom{r-1}{r-1} + \binom{r}{r-1}(qs) + \binom{r+1}{r-1}(qs)^2 + \dots$$

$$= (ps)^r [1 + r(qs) + \frac{r(r+1)}{2}(qs)^2 + \dots$$

$$= (ps)^r [1 + \binom{-r}{1}(-qs) + \binom{-r}{2}(-qs)^2 + \dots$$

$$= (ps)^r \sum_{k=r}^{\infty} \binom{-r}{k}(-qs)^k$$

$$= \frac{(ps)^r}{(1-qs)^r}$$

Hence

$$H(s) = [G(s)]^N$$

$$= [\frac{ps}{1-qs}]^{Nr}$$

which is the pgf of a Negative Binomial distribution with parameters Nr and p.

Thus the distribution of the sum of Negative Binomial independent and identically distributed random variables is Negative Binomial.

2.5.3.5 The sum of Poisson variables

Given

$$p_k = Prob(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \quad ; k = 0, 1, 2, \dots$$

then

$$
\begin{aligned}
G(s) &= \sum_{k=0}^{\infty} p_k s^k = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}.s^k \\
&= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} \\
&= e^{-\lambda}[1 + \frac{(\lambda s)}{1!} + \frac{(\lambda s)^2}{2!} + \ldots \\
&= e^{-\lambda}.e^{\lambda s} \\
&= e^{-\lambda(1-s)}
\end{aligned}
$$

Thus

$$
\begin{aligned}
H(s) &= [G(s)]^N \\
&= [e^{-\lambda(1-s)}]^N \\
&= e^{-N\lambda(1-s)}
\end{aligned}
$$

which is the pgf of a Poisson distribution with parameter Nλ.
Thus the distribution of the sum of Poisson independent and identically distributed random variables is Poisson.

2.5.3.6 The sum of Uniform variables

Given

$$
p_k = Prob(X = k) = \frac{1}{a} \ ; k = 1, 2, \ldots, a
$$

then

$$
\begin{aligned}
G(s) &= \sum_{k=1}^{a} p_k s^k = \sum_{k=1}^{a} \frac{1}{a} s^k \\
&= \frac{1}{a} \sum_{k=1}^{a} s^k \\
&= \frac{1}{a}[s + s^2 + s^3 + \ldots + s^a] \\
&= \frac{s}{a}[1 + s + s^2 + \ldots + s^{a-1}] \\
&= \frac{s}{a}[\frac{1 - s^a}{1 - s}]
\end{aligned}
$$

Thus

$$\begin{aligned} H(s) &= [G(s)]^N \\ &= [\frac{s}{a}[\frac{1-s^a}{1-s}]]^N \\ &= \frac{s^N}{a^N}[\frac{1-s^a}{1-s}]^N \end{aligned}$$

which is a pgf, but not of uniform distribution.

2.5.3.7 The sum of Logarithmic variables

Given

$$p_k = Prob(X = k) = -\frac{p^k}{klog(1-p)} \quad ; k = 1, 2, \ldots$$

then

$$\begin{aligned} G(s) &= \sum_{k=1}^{\infty} p_k s^k = \sum_{k=1}^{\infty} -\frac{p^k}{klog(1-p)} s^k \\ &= \sum_{k=1}^{\infty} -\frac{(ps)^k}{klog(1-p)} \\ &= -\frac{1}{log(1-p)} \sum_{k=1}^{\infty} \frac{(ps)^k}{k} \\ &= -\frac{1}{log(1-p)}[ps + \frac{(ps)^2}{2} + \frac{(ps)^3}{3} + \ldots \\ &= -\frac{1}{log(1-p)} - log(1-ps) \\ &= \frac{log(1-ps)}{log(1-p)} \end{aligned}$$

Thus

$$\begin{aligned} H(s) &= [G(s)]^N \\ &= [\frac{log(1-ps)}{log(1-p)}]^N \end{aligned}$$

which is a pgf, but not of a logarithmic distribution.

2.6 Conditional probability distributions given the sum of two independent random variables.

We can use sums of iid random variables to obtain conditional distribution functions.

2.6.1 Conditional probability distribution given sum of two Binomial random variables.

Let X and Y be two independent random variables from a binomial distribution that is X $\sim$ B(m,p) and Y $\sim$ B(n,p).Then

$$
\begin{aligned}
Prob(X = x/X + Y = x + y) &= \frac{Prob(X = x, X + Y = x + y)}{Prob(X + Y = x + y)} \\
&= \frac{Prob(X = x)Prob(Y = y)}{Prob(X + Y = x + y)} \\
&= \frac{\binom{m}{x}p^x(1-p)^{m-x}\binom{n}{y}p^y(1-p)^{n-y}}{(m+n)(x+y)p^{x+y}(1-p)^{(m+n)-(x+y)}} \\
&= \frac{\binom{m}{x}\binom{n}{y}}{\binom{m+n}{x+y}} \qquad x = 0, 1, 2, \ldots, m
\end{aligned}
$$

Let $r = x + y$,then

$$
Prob(X = x/X + Y = r) = \frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}} \qquad x = 0, 1, 2, \ldots, r \quad (2.32)
$$

where $r \leq m + n$
which is a Hyper-geometric distribution.

2.6.2 Conditional probability distribution given sum of two Poisson random variables.

Let X and Y be two independent random variables from a Poison distribution that is X $\sim$ Poisson(λ) and Y $\sim$ Poisson(μ).Then

$$Prob(X = x/X + Y = x + y) = \frac{Prob(X = x, X + Y = x + y)}{Prob(X + Y = x + y)}$$
$$= \frac{Prob(X = x)Prob(Y = y)}{Prob(X + Y = x + y)}$$
$$= \frac{\frac{\lambda^x e^{-\lambda}}{x!}.\frac{\mu^y e^{-\mu}}{y!}}{(\lambda + \mu)^{x+y} e^{-(\lambda+\mu)}(x + y)!}$$
$$= \frac{\lambda^x \mu^y (x + y)!}{x!y!(\lambda + \mu)^{x+y}}$$

Let $x + y = n$ and $x = k$
Thus

$$Prob(X = x/X + Y = x + y) = \frac{\lambda^k \mu^{n-k} (n)!}{k!(n - k)!(\lambda + \mu)^n}$$

$$= \binom{n}{k}(\frac{\lambda}{\lambda + \mu})^k(\frac{\mu}{\lambda + \mu})^{n-k} \qquad k = 0, 1, 2, \ldots, n \qquad (2.33)$$

which is a binomial distribution with parameters n and $\frac{\lambda}{\lambda+\mu}$.

2.6.3 Conditional probability distribution given sum of two Geometric random variables.

Let X and Y be two independent random variables from a Geometric distribution that is X $\sim$ Geometric(p) and Y $\sim$ Geometric(p).Then

$$Prob(X = x/X + Y = x + y) = \frac{Prob(X = x, X + Y = x + y)}{Prob(X + Y = x + y)}$$
$$= \frac{Prob(X = x)Prob(Y = y)}{Prob(X + Y = x + y)}$$

due of independence.
Let $x + y = n \Longrightarrow y = n - x$

Therefore

$$Prob(X = x/X + Y = n) = \frac{(1-p)^x p(1-p)^{n-x} p}{\binom{2+n-1}{n}(1-p)^n p^2}$$
$$= \frac{1}{\binom{n+1}{n}}$$

$$= \frac{1}{n+1} \qquad x = 0, 1, 2, \ldots, n \tag{2.34}$$

which is the Uniform distribution.

2.7 Compound distributions based on varying parameters of probability mass functions.

2.7.1 Compound distributions from Binomial Distribution

Let
$p_k = Prob(X = k) = \binom{N}{k} p^k q^{N-k} \qquad k = 0, 1, 2, \ldots, N.$ where $q = 1 - p$ which is Binomial with parameters N and p.

2.7.1.1 Beta-Binomial Distribution

Let p $\sim$ Beta(α,β) i.e. the parameter p is from the Beta distribution which is a continous probability distribution; i.e.,

$$Prob(P = p) = \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1}(1-p)^{\beta-1}; 0 < p < 1$$

$$
\begin{aligned}
p_k &= \int Prob(X = k/P = p)Prob(P = p)dp \\
&= \int_0^1 \binom{n}{k} p^k (1-p)^{n-k} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1}(1-p)^{\beta-1} dp \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \binom{n}{k} \int_0^1 p^{\alpha+k-1}(1-p)^{n+\beta-k-1} dp \\
&= \binom{n}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} B(\alpha+k, n+\beta-k)
\end{aligned}
$$

Thus

$$
p_k = \binom{n}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \frac{\Gamma(\alpha+k)\Gamma(\beta+n-k)}{\Gamma(\alpha+\beta+n)} \quad k = 0, 1, 2, \ldots, n \quad (2.35)
$$

If $\alpha = \beta = 1$,then

$$
p_k = \binom{n}{k} \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)}
$$

$$
= \frac{1}{n+1} \qquad k = 0, 1, 2, \ldots, n \qquad (2.36)
$$

which is the Uniform distribution.

2.7.1.2 Binomial-Binomial Distribution

Let N $\sim$ Binomial(M,q) i.e.

$$
p_n = Prob(N = n) = \binom{M}{n} \theta^n (1-\theta)^{M-n} \qquad n = 0, 1, 2, \ldots, M
$$

Therefore

$$
\begin{aligned}
p_k = Prob(X = k) &= \sum_{n=0}^{M} Prob(X = k/N = n)Prob(N = n) \\
&= \sum_{n=0}^{M} \binom{n}{k} p^k (1-p)^{n-k} \binom{M}{n} \theta^n (1-\theta)^{M-n} \\
&= \sum_{n=0}^{M} \binom{n}{k} \binom{M}{n} p^k \theta^n (1-p)^{n-k} (1-\theta)^{M-n} \\
&= (\theta p)^k \sum_{n=0}^{M} \binom{n}{k} \binom{M}{n} \theta^{n-k} (1-p)^{n-k} (1-\theta)^{M-n} \\
&= (\theta p)^k \sum_{n=0}^{M} \binom{n}{k} \binom{M}{n} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= (\theta p)^k \sum_{n=0}^{M} \frac{n!}{k!(n-k)!} \frac{M!}{n!(M-n)!} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= (\theta p)^k \sum_{n=0}^{M} \frac{M!}{k!(n-k)!(M-n)!} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= (\theta p)^k \sum_{n=0}^{M} \frac{M!(M-k)!}{k!(M-k)!(n-k)!(M-n)!} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= (\theta p)^k \frac{M!}{k!(M-k)!} \sum_{n=0}^{M} \frac{(M-k)!}{(n-k)!(M-n)!} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= (\theta p)^k \binom{M}{k} \sum_{n=0}^{M} \binom{M-k}{n-k} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= \binom{M}{k} (\theta p)^k [(1-\theta) + \theta(1-p)]^{M-k}
\end{aligned}
$$

Thus

$$
p_k = \binom{M}{k} (\theta p)^k (1-\theta p)^{M-k} \qquad k = 0, 1, 2, \ldots, M \qquad (2.37)
$$

which is a binomial distribution with parameters θp and M.

Using the pgf technique i.e.

$$
\begin{aligned}
G(s) &= \sum_{k=0}^{M} p_k s^k \\
&= \sum_{k=0}^{M}\sum_{n=0}^{M} \binom{n}{k} p^k (1-p)^{n-k} \binom{M}{n} \theta^n (1-\theta)^{M-n} s^k \\
&= \sum_{k=0}^{M} (ps)^k \sum_{n=0}^{M} \binom{n}{k} (1-p)^{n-k} \binom{M}{n} \theta^{n-k} \theta^k (1-\theta)^{M-n} \\
&= \sum_{k=0}^{M} (\theta ps)^k \sum_{n=0}^{M} \binom{n}{k} [\theta(1-p)]^{n-k} \binom{M}{n} (1-\theta)^{M-n} \\
&= \sum_{k=0}^{M} (\theta ps)^k \sum_{n=0}^{M} \frac{n!}{k!(n-k)!} [\theta(1-p)]^{n-k} \frac{M!}{n!(M-n)!} (1-\theta)^{M-n} \\
&= \sum_{k=0}^{M} (\theta ps)^k \sum_{n=0}^{M} \frac{M!(M-k)!}{k!(n-k)!(M-k)!(M-n)!} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= \sum_{k=0}^{M} \frac{M!}{k!(M-k)!} (\theta ps)^k \sum_{n=0}^{M} \frac{(M-k)!}{(n-k)!(M-n)!} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= \sum_{k=0}^{M} \binom{M}{k} (\theta ps)^k \sum_{n=0}^{M} \binom{M-k}{n-k} [\theta(1-p)]^{n-k} (1-\theta)^{M-n} \\
&= \sum_{k=0}^{M} \binom{M}{k} (\theta ps)^k [(1-\theta) + \theta(1-p)]^{M-k} \\
&= \sum_{k=0}^{M} \binom{M}{k} (\theta ps)^k [1-\theta p]^{M-k} \\
&= [(1-\theta p) + \theta ps]^M
\end{aligned}
$$

Hence

$$
G(s) = [(1-p\theta) + \theta ps]^M \qquad (2.38)
$$

which is the pgf of a binomial distribution with parameters θp and M. Thus the Binomial-Binomial Distribution is Binomial.

2.7.1.3 Poisson-Binomial Distribution

Let N ~ Poisson(λ) i.e.

$$p_n = Prob(N = n) = \frac{e^{-\lambda}\lambda^n}{n!} \qquad n = 0, 1, 2, \ldots$$

Therefore

$$\begin{aligned}
p_k = Prob(X = k) &= \sum_{n=0}^{\infty} Prob(X = k/N = n) Prob(N = n) \\
&= \sum_{n=0}^{\infty} \binom{n}{k} p^k q^{n-k} \frac{e^{-\lambda}\lambda^n}{n!} \\
&= e^{-\lambda} p^k \sum_{n=0}^{\infty} \binom{n}{k} q^{n-k} \frac{\lambda^n}{n!} \\
&= e^{-\lambda} (\lambda p)^k \sum_{n=0}^{\infty} \binom{n}{k} \frac{(\lambda q)^{n-k}}{n!} \\
&= e^{-\lambda} (\lambda p)^k \sum_{n=0}^{\infty} \frac{n!}{k!(n-k)!} \frac{(\lambda q)^{n-k}}{n!} \\
&= \frac{e^{-\lambda}(\lambda p)^k}{k!} \sum_{n=k}^{\infty} \frac{(\lambda q)^{n-k}}{(n-k)!} \\
&= \frac{e^{-\lambda}(\lambda p)^k}{k!} e^{\lambda q} \\
&= \frac{e^{-\lambda}(\lambda p)^k}{k!} e^{\lambda(1-p)}
\end{aligned}$$

Thus

$$p_k = \frac{e^{-\lambda p}(\lambda p)^k}{k!} \qquad k = 0, 1, 2, \ldots \tag{2.39}$$

which is the Poisson distribution with parameter λp.

Using the pgf technique i.e.

$$\begin{aligned}
G(s) &= \sum_{k=0}^{\infty} p_k s^k \\
&= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \binom{n}{k} p^k q^{n-k} \frac{e^{-\lambda}\lambda^n}{n!} s^k \\
&= e^{-\lambda}\sum_{k=0}^{\infty}(ps)^k \sum_{n=0}^{\infty} \binom{n}{k} q^{n-k}\frac{\lambda^n}{n!} \\
&= e^{-\lambda}\sum_{k=0}^{\infty}(ps)^k \sum_{n=0}^{\infty} \binom{n}{k} \frac{(\lambda q)^{n-k}}{n!}\lambda^k \\
&= e^{-\lambda}\sum_{k=0}^{\infty}(\lambda ps)^k \sum_{n=0}^{\infty} \frac{n!}{k!(n-k)!}\frac{(\lambda q)^{n-k}}{n!} \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda ps)^k}{k!} \sum_{n=0}^{\infty} \frac{(\lambda q)^{n-k}}{(n-k)!} \\
&= e^{-\lambda}e^{\lambda ps}e^{\lambda q} \\
&= e^{-\lambda(1-(1-p)-ps)} \\
&= e^{-\lambda p(1-s)}
\end{aligned}$$

Thus

$$G(s) = e^{-\lambda p(1-s)} \tag{2.40}$$

which is the pgf of a Poisson distribution with parameter λp.
Thus the Poisson-Binomial Distribution is Poisson.

2.7.1.4 Logarithmic-Binomial Distribution

Let N $\sim$ Logarithmic(α) i.e.

$$p_n = Prob(N=n) = \frac{\alpha^n}{-nlog(1-\alpha)} \qquad n = 1, 2, \ldots$$

Let $c = \frac{1}{-log(1-\alpha)}$ then

$$Prob(N=n) = \frac{c\alpha^n}{n}$$

Therefore

$$
\begin{aligned}
p_k = Prob(X = k) &= \sum_{n=1}^{\infty} Prob(X = k/N = n)Prob(N = n) \\
&= \sum_{n=1}^{\infty} \tbinom{n}{k} p^k (1-p)^{n-k} \frac{c\alpha^n}{n} \\
&= c(\frac{p}{1-p})^k \sum_{n=1}^{\infty} \tbinom{n}{k} \frac{[\alpha(1-p)]^n}{n} \\
&= c(\frac{p}{1-p})^k \sum_{n=k}^{\infty} \frac{n!}{k!(n-k)!} \frac{[\alpha(1-p)]^n}{n} \\
&= c(\frac{p}{1-p})^k \sum_{n=k}^{\infty} \frac{(n-1)!}{k!(n-k)!} \frac{[\alpha(1-p)]^n}{n} \\
&= \frac{c}{k}(\frac{p}{1-p})^k \sum_{n=k}^{\infty} \frac{(n-1)!}{(k-1)!(n-k)!} \frac{[\alpha(1-p)]^n}{n} \\
&= \frac{c}{k}(\frac{p}{1-p})^k \sum_{n=k}^{\infty} \tbinom{n-1}{k-1} \frac{[\alpha(1-p)]^n}{n} \\
&= \frac{c}{k}(\frac{p}{1-p})^k [\tbinom{k-1}{k-1}[\alpha(1-p)]^k + \tbinom{k}{k-1}[\alpha(1-p)]^{k+1} + \dots \\
&= \frac{c}{k}(\frac{p}{1-p})^k [\alpha(1-p)]^k [1 + k[\alpha(1-p)] + \frac{k(k+1}{2}[\alpha(1-p)]^2 + \dots \\
&= \frac{c}{k}(\alpha p)^k [1 + \tbinom{k}{1}[\alpha(1-p)] + \tbinom{k+1}{2}[\alpha(1-p)]^2 + \dots \\
&= \frac{c}{k}(\alpha p)^k \sum_{r=0}^{\infty} \tbinom{k+r+1}{r}[\alpha(1-p)]^r \\
&= \frac{c}{k}(\alpha p)^k \sum_{r=0}^{\infty} (-1)^r \tbinom{-k}{r}[\alpha(1-p)]^r \\
&= \frac{c}{k}(\alpha p)^k [1 - \alpha(1-p)]^{-k} \\
&= \frac{c}{k} \frac{(\alpha p)^k}{[1 - \alpha(1-p)]^k} \\
&= \frac{c}{k}[\frac{(\alpha p)}{1 - \alpha(1-p)}]^k
\end{aligned}
$$

Let

$$\beta = \frac{(\alpha p)}{1-\alpha(1-p)}$$

Hence

$$p_k = \frac{c}{k}\beta^k = \frac{\beta^k}{-klog(1-\alpha)} \qquad k = 1, 2, \ldots$$

But

$$p_0 + \sum_{k=1}^{\infty} p_k = 1$$

$$\Longrightarrow p_0 + \sum_{k=1}^{\infty} \frac{c}{k}\beta^k = 1$$

$$\Longrightarrow p_0 + c[-log(1-\beta)] = 1$$

Therefore

$$\begin{aligned} p_0 &= 1 + clog(1-\beta) \\ &= \frac{log(1-\beta)}{log(1-\alpha)} \end{aligned}$$

But

$$\begin{aligned} 1-\beta = 1 - \frac{\alpha p}{1-\alpha+\alpha p} \\ = \frac{1-\alpha}{1-\alpha+\alpha p} \end{aligned}$$

Therefore

$$\begin{aligned} p_0 &= 1 - \frac{log(\frac{1-\alpha}{1-\alpha+\alpha p}}{log(1-\alpha)} \\ &= \frac{log(1-\alpha) - log(\frac{1-\alpha}{1-\alpha+\alpha p}}{log(1-\alpha)} \\ &= \frac{log(1-\alpha)\frac{1-\alpha}{1-\alpha+\alpha p}}{log(1-\alpha)} \\ &= \frac{1-\alpha+\alpha p}{log(1-\alpha)} \end{aligned}$$

Therefore

$$p_k = \begin{cases} \frac{1-\alpha+\alpha p}{log(1-\alpha)} & for \;\; k=0 \\ \frac{(\frac{\alpha p}{1-\alpha+\alpha p})^k}{klog(1-\alpha)} & for \;\; k=1,2,3,... \end{cases} \tag{2.41}$$

Using pgf technique,we have

$$\begin{aligned} G(s) &= \sum_{k=0}^{n}\sum_{n=1}^{\infty}\binom{n}{k}(ps)^k(1-p)^{n-k}\frac{c\alpha^n}{n} \\ &= \sum_{n=1}^{\infty}[(1-p)+ps]^n\frac{c\alpha^n}{n} \\ &= c\sum_{n=1}^{\infty}\frac{[\alpha(1-p)+\alpha ps]^n}{n} \\ &= c[-log(1-\alpha(1-p)-\alpha ps)] \\ &= \frac{log[1-\alpha(1-p)-\alpha ps]}{log(1-\alpha)} \\ &= \frac{log[1-\alpha(1-p)][1-\frac{\alpha p}{1-\alpha+\alpha p}s]}{log(1-\alpha)} \\ &= \frac{log(1-\alpha+\alpha p)+log(1-\frac{\alpha p}{1-\alpha+\alpha p}s)}{log(1-\alpha)} \\ &= \frac{log(1-\alpha+\alpha p)}{log(1-\alpha)} - \sum_{k=1}^{\infty}[\frac{(\frac{\alpha p}{1-\alpha+\alpha p})^k s^k}{klog(1-\alpha)}] \end{aligned}$$

which is a pgf.
Therefore

$$p_k = \begin{cases} \frac{1-\alpha+\alpha p}{log(1-\alpha)} & for \;\; k=0 \\ \frac{(\frac{\alpha p}{1-\alpha+\alpha p})^k}{klog(1-\alpha)} & for \;\; k=1,2,3,... \end{cases} \tag{2.42}$$

2.7.2 Compound distributions from Poisson Distribution

The Poisson distribution is an example of a discrete probability distribution with

$$p_k = Prob(X=k) = \frac{e^{-\lambda}\lambda^k}{k!} \qquad k=0,1,2,\ldots$$

λ is the only parameter.

2.7.2.1 Exponential-Poisson Distribution

Let $\lambda \sim$ Exponential($\frac{1}{\mu}$) i.e.

$$p_\lambda = Prob(\Lambda = \lambda) = \frac{1}{\mu}e^{-\frac{\lambda}{\mu}} \qquad \lambda > 0$$

$$\begin{aligned} p_k &= \int Prob(X = k/\Lambda = \lambda)Prob(\Lambda = \lambda)d\lambda \\ &= \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!}.\frac{1}{\mu}e^{-\frac{\lambda}{\mu}}d\lambda \\ &= \frac{1}{k!\mu}\int_0^\infty \lambda^k e^{-(\frac{\mu+1}{\mu})\lambda}d\lambda \end{aligned}$$

Let

$$y = (\frac{\mu+1}{\mu})\lambda$$

$$\Longrightarrow \lambda = \frac{\mu}{\mu+1}y$$

and

$$d\lambda = \frac{\mu}{\mu+1}dy$$

Therefore

$$\begin{aligned} p_k &= \frac{1}{k!\mu}\int_0^\infty [\frac{\mu}{\mu+1}y]^k e^{-y}\frac{\mu}{\mu+1}dy \\ &= \frac{1}{k!\mu}[\frac{\mu}{\mu+1}]^{k+1}\int_0^\infty y^k e^{-y}dy \\ &= \frac{1}{k!\mu}[\frac{\mu}{\mu+1}]^{k+1}\Gamma(k+1) \\ &= \frac{1}{\mu}[\frac{\mu}{\mu+1}]^{k+1} \end{aligned}$$

$$= \quad (\frac{1}{\mu+1})(\frac{\mu}{\mu+1})^k \qquad k = 0, 1, 2, \ldots \tag{2.43}$$

which is the Geometric distribution where $p = \frac{1}{\mu+1}$ and $q = 1 - p = \frac{\mu}{\mu+1}$

Using the pgf technique we have

$$\begin{aligned}
G(s) &= \sum_{k=0}^{\infty} p_k s^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!\mu} \int_0^{\infty} \lambda^k e^{-(\frac{\mu+1}{\mu})\lambda} d\lambda s^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!\mu} (\frac{\mu}{\mu+1})^{k+1} \Gamma(k+1) s^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} (\frac{(\mu)^k}{(\mu+1)^{k+1}} s^k k! \\
&= \sum_{k=0}^{\infty} \frac{(\mu)^k}{(\mu+1)^{k+1}} s^k \\
&= \frac{1}{\mu+1} \sum_{k=0}^{\infty} (\frac{\mu}{\mu+1} s)^k \\
&= p \sum_{k=0}^{\infty} (qs)^k \\
&= \frac{p}{1-qs}
\end{aligned}$$

which is the pgf of a Geometric distribution with k being the number of failures before a success.
Thus the Exponential-Poisson Distribution is Geometric.

2.7.2.2 Gamma-Poisson Distribution

Let $\lambda \sim$ Gamma(α,β) i.e.

$$p_\lambda = Prob(\Lambda = \lambda) = \frac{\beta^\alpha}{\Gamma\alpha} \lambda^{\alpha-1} e^{-\beta\lambda} \quad \lambda > 0$$

$$p_k = \int Prob(X = k/\Lambda = \lambda).Prob(\Lambda = \lambda)d\lambda$$
$$= \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!}\frac{\beta^\alpha}{\Gamma\alpha}\lambda^{\alpha-1}e^{-\beta\lambda}d\lambda$$
$$= \frac{\beta^\alpha}{k!\Gamma\alpha}\int_0^\infty \lambda^{\alpha+k-1}e^{-(\beta+1)\lambda}d\lambda$$

Let

$$y = (\beta + 1)\lambda$$
$$\Longrightarrow \lambda = \frac{y}{\beta + 1}$$

and

$$d\lambda = \frac{dy}{\beta + 1}$$

Therefore

$$p_k = \frac{\beta^\alpha}{k!\Gamma\alpha}\int_0^\infty \int_0^\infty (\frac{y}{\beta+1})^{\alpha+k-1}e^{-y}\frac{dy}{\beta+1}$$
$$= \frac{\beta^\alpha}{k!\Gamma\alpha}(\frac{1}{\beta+1})^{\alpha+k}\int_0^\infty (\frac{y}{\beta+1})^{\alpha+k-1}e^{-y}dy$$
$$= \frac{\beta^\alpha}{k!\Gamma\alpha}\frac{1}{(\beta+1)^{\alpha+k}}\Gamma(\alpha+k)$$
$$= \frac{(\alpha+k-1)!}{k!(\alpha-1)!}(\frac{1}{\beta+1})^k(\frac{\beta}{\beta+1})^\alpha$$

$$= \binom{\alpha+k-1}{k}(\frac{1}{\beta+1})^k(\frac{\beta}{\beta+1})^\alpha \qquad k = 0, 1, 2, \ldots \qquad (2.44)$$

which is the Negative Binomial distribution with parameters α and $\frac{\beta}{\beta+1}$.

Using the pgf technique we have

$$\begin{aligned}
G(s) &= \sum_{k=0}^{\infty} p_k s^k \\
&= \sum_{k=0}^{\infty} \int_0^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \frac{\beta^\alpha}{\Gamma\alpha} \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda s^k \\
&= \sum_{k=0}^{\infty} \frac{\beta^\alpha}{k!\Gamma\alpha} \int_0^{\infty} \lambda^{\alpha+k-1} e^{-(\beta+1)\lambda} d\lambda s^k \\
&= \sum_{k=0}^{\infty} \binom{\alpha+k-1}{k} (\frac{\beta}{\beta+1})^\alpha (\frac{1}{\beta+1})^k s^k \\
&= (\frac{\beta}{\beta+1})^\alpha \sum_{k=0}^{\infty} \binom{\alpha+k-1}{k} (\frac{s}{\beta+1})^k \\
&= (\frac{\beta}{\beta+1})^\alpha \sum_{k=0}^{\infty} (-1)^k \binom{-\alpha}{k} (\frac{s}{\beta+1})^k \\
&= (\frac{\beta}{\beta+1})^\alpha (1 - \frac{1}{\beta+1} s)^{-\alpha}
\end{aligned}$$

which is the pgf of a Negative Binomial distribution with parameters α and $\frac{\beta}{\beta+1}$

Thus the Gamma-Poisson Distribution is Negative Binomial.

2.7.2.3 Logarithmic-Poisson Distribution

Let $\lambda \sim$ Logarithmic(p) i.e.

$$p_\lambda = Prob(\Lambda = \lambda) = \frac{p^\lambda}{-\lambda log(1-p)} \qquad \lambda = 1, 2, \ldots$$

$$\begin{aligned}
p_k &= \sum_{\lambda=1}^{\infty} Prob(X = k/\Lambda = \lambda) Prob(\Lambda = \lambda) \\
&= \sum_{\lambda=1}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \cdot \frac{p^\lambda}{-\lambda log(1-p)} \qquad k = 0, 1, 2, \ldots
\end{aligned}$$

which cannot be expressed explicitly.

Using the pgf technique,we have

$$
\begin{aligned}
G(s) &= \sum_{k=0}^{\infty} p_k s^k \\
&= \sum_{k=0}^{\infty}\sum_{\lambda=1}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}\frac{p^\lambda}{-\lambda log(1-p)} s^k \\
&= \sum_{k=0}^{\infty}\sum_{\lambda=1}^{\infty} \frac{e^{-\lambda}(\lambda s)^k}{k!}\frac{p^\lambda}{-\lambda log(1-p)} \\
&= \frac{1}{-log(1-p)}\sum_{\lambda=1}^{\infty}\frac{e^{-\lambda}p^\lambda}{\lambda}\sum_{k=0}^{\infty}\frac{(\lambda s)^k}{k!} \\
&= \frac{1}{-log(1-p)}\sum_{\lambda=1}^{\infty}\frac{e^{-\lambda}p^\lambda}{\lambda}e^{\lambda s} \\
&= \frac{1}{-log(1-p)}\sum_{\lambda=1}^{\infty}\frac{e^{-\lambda(1-s)}p^\lambda}{\lambda} \\
&= \frac{1}{-log(1-p)}[\frac{e^{-(1-s)}p}{1}+\frac{e^{-2(1-s)}p}{2}+\frac{e^{-3(1-s)}p}{3}+\ldots] \\
&= \frac{1}{-log(1-p)}[\frac{e^{-(1-s)}p}{1}+\frac{(e^{-(1-s)}p)^2}{2}+\frac{(e^{-(1-s)}p)^3}{3}+\ldots] \\
&= \frac{1}{-log(1-p)} \times -log[1-pe^{-(1-s)}] \\
&= \frac{log[1-pe^{-(1-s)}]}{log(1-p)}
\end{aligned}
$$

2.7.2.4 Poisson-Poisson Distribution

Let $\lambda \sim$ Poisson(μ) i.e.

$$p_\lambda = Prob(\Lambda = \lambda) = \frac{e^{-\mu}\mu^\lambda}{\lambda!}$$

$$p_k = \sum_{\lambda=0}^{\infty} Prob(X = k/\Lambda = \lambda).Prob(\Lambda = \lambda)$$

$$= \sum_{\lambda=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}.\frac{e^{-\mu}\mu^\lambda}{\lambda!} \qquad k = 0, 1, 2, \ldots$$

which cannot be expressed explicitly.

Using the pgf technique,we have

$$G(s) = \sum_{k=0}^{\infty} p_k s^k$$

$$= \sum_{k=0}^{\infty}\sum_{\lambda=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}.\frac{e^{-\mu}\mu^\lambda}{\lambda!}s^k$$

$$= \sum_{\lambda=0}^{\infty} \frac{e^{-\lambda}\mu^\lambda e^{-\mu}}{\lambda!} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!}$$

$$= \sum_{\lambda=0}^{\infty} \frac{e^{-(\lambda+\mu)}\mu^\lambda}{\lambda!} e^{\lambda s}$$

$$= \sum_{\lambda=0}^{\infty} \frac{e^{-\lambda-\mu+\lambda s}\mu^\lambda}{\lambda!}$$

$$= e^{-\mu} \sum_{\lambda=0}^{\infty} \frac{e^{-\lambda(1-s)}\mu^\lambda}{\lambda!}$$

$$= e^{-\mu} \sum_{\lambda=0}^{\infty} \frac{[e^{-(1-s)}\mu]^\lambda}{\lambda!}$$

$$= e^{-\mu} exp[\mu e^{-(1-s)}]$$

$$= e^{-\mu+\mu e^{-(1-s)}}$$

$$= e^{-\mu[1-e^{-(1-s)}]} \qquad k = 0, 1, 2, \ldots \qquad (2.45)$$

It is not easy to express p_k,the coefficient of s^k in G(s).

2.7.3 Compound distributions from Geometric Distribution

The Geometric distribution is an example of a discrete probability distribution with

<u>Case 1</u>

Let X be the number of failures before the first success,then from equation (2.16)

$$p_k = Prob(X = k) = q^k p \qquad k = 0, 1, 2, \ldots$$

where $q = 1 - p$
p is the only parameter.

<u>Case 2</u>

Let X be the total number of trials required to achieve the first success,then from equation (2.17)

$$p_k = Prob(X = k) = q^{k-1} p \qquad k = 1, 2, \ldots$$

where $q = 1 - p$
p is the only parameter.

2.7.3.1 Beta-Geometric Distribution

<u>**Type I**</u>

Let p $\sim$ Beta(α,β) i.e. the parameter p is from the Beta distribution which is a continous probability distribution.

$$p_p = Prob(P = p) = \frac{\Gamma(\alpha + \beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1} q^{\beta-1} \quad ; o < p < 1$$

where $q = 1 - p$

$$
\begin{aligned}
p_k &= \int Prob(x = k/P = p) Prob(P = p) dp \\
&= \int_0^1 q^k p \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1} q^{\beta-1} dp \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \int_0^1 p^{\alpha} q^{\beta+k-1} dp \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} B(\alpha+1, \beta+k) \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \frac{\Gamma(\alpha+1)\Gamma(\beta+k)}{\Gamma(\alpha+\beta+k+1)}
\end{aligned}
$$

$$
= \frac{\alpha\Gamma(\alpha+\beta)\Gamma(\beta+k)}{\Gamma\beta\Gamma(\alpha+\beta+k+1)} \qquad k = 0, 1, 2, \ldots \tag{2.46}
$$

which is a probability mass function.

Type II

Let p $\sim$ Beta(α,β) i.e. the parameter p is from the Beta distribution which is a continous probability distribution.

$$
p_p = Prob(P = p) = \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1} q^{\beta-1}
$$

where $q = 1 - p$

$$
\begin{aligned}
p_k &= \int Prob(x = k/P = p) Prob(P = p) dp \\
&= \int_0^1 q^{k-1} p \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1} q^{\beta-1} dp \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \int_0^1 p^{(\alpha+1)-1} q^{(\beta+k-1)-1} dp \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} B(\alpha+1, \beta+k-1) \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \frac{\Gamma(\alpha+1)\Gamma(\beta+k-1)}{\Gamma(\alpha+\beta+k)}
\end{aligned}
$$

$$= \frac{\alpha\Gamma(\alpha+\beta)\Gamma(\beta+k-1)}{\Gamma\beta\Gamma(\alpha+\beta+k)} \qquad k = 1, 2, \ldots \qquad (2.47)$$

which is a probability mass function.

2.7.4 Compound distributions from Negative Binomial Distribution

2.7.4.1 Beta-Negative Binomial Distribution

Type I

Let X be the number of failures before the r^{th} success,then from equation (2.11)

$$p_k = Prob(X = k) = \binom{r+k-1}{k} q^k p^r \qquad k = 0, 1, 2, \ldots$$

where $q = 1 - p$,p and r are parameters.
Let p $\sim$ Beta(α,β) i.e. the parameter p is from the Beta distribution which is a continous probability distribution.

$$p_p = Prob(P = p) = \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1} q^{\beta-1}$$

where $q = 1 - p$

$$\begin{aligned} p_k &= \int_0^1 Prob(x = k/P = p).Prob(P = p)dp \\ &= \int_0^1 \binom{r+k-1}{k} q^k p^r \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1} q^{\beta-1} dp \\ &= \binom{r+k-1}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \int_0^1 p^{\alpha+r-1} q^{\beta+k-1} dp \\ &= \binom{r+k-1}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} B(\alpha+r, \beta+k) \end{aligned}$$

Thus

$$p_k = \binom{r+k-1}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \frac{\Gamma(\alpha+r)\Gamma(\beta+k)}{\Gamma(\alpha+\beta+k+r)} \qquad k = 0, 1, 2, \ldots \qquad (2.48)$$

which is a probability mass function.

Type II

Let X be the total number of trials required to achieve the r^{th} success,then from equation (2.13)

$$p_k = Prob(X = k) = \binom{k-1}{k-r} q^{k-r} p^r = \binom{k-1}{r-1} q^{k-r} p^r \qquad k = r, r+1, r+2, \ldots$$

where $q = 1 - p$,p and r are parameters.
Let p ~ Beta(α,β) i.e. the parameter p is from the Beta distribution which is a continous probability distribution.

$$p_p = Prob(P = p) = \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1} q^{\beta-1}$$

where $q = 1 - p$

$$\begin{aligned} p_k &= \int_0^1 Prob(x = k/P = p) Prob(P = p) dp \\ &= \int_0^1 \binom{k-1}{r-1} q^{k-r} p^r \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} p^{\alpha-1} q^{\beta-1} \\ &= \binom{k-1}{r-1} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \int_0^1 p^{\alpha+r-1} q^{\beta+k-r-1} dp \\ &= \binom{k-1}{r-1} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} B(\alpha+r, \beta+k-r) \end{aligned}$$

Thus

$$p_k = \binom{k-1}{r-1} \frac{\Gamma(\alpha+\beta)}{\Gamma\alpha\Gamma\beta} \frac{\Gamma(\alpha+r)\Gamma(\beta+k-r)}{\Gamma(\alpha+\beta+k)} \qquad (2.49)$$

for $k = r, r+1, r+2, \ldots$
which is a probability mass function.

2.8 Compound distributions based on the sum of a random number of independent and identically distributed random variables.

Let $S_N = X_1 + X_2 + \ldots + X_N$ where the $X_i's$ are iid random variables.N is also a random variable independent of the $X_i's$.Let

$G(s) = E(s^X)$ the pgf of X
$F(s) = E(s^N)$ the pgf of N
$H(s) = E(s^{S_N})$ the pgf of S_N
Therefore

$$\begin{aligned} H(s) &= E[s^{S_N}] \\ &= EE[s^{S_N}/N] \\ &= EE[s^{X_1+X_2+\ldots+X_N}] \\ &= E[E(s^{X_1}) + E(s^{X_2}) + \ldots + E(s^{X_N})] \\ &= E[G(s)]^N \end{aligned}$$

Thus

$$H(s) = F_N[G_X(s)] \tag{2.50}$$

which is a compound distribution.

2.8.1 Compound Distributions for Standard Discrete Distributions

2.8.1.1 Compound Binomial Distribution

From $H(s) = F_N[G_X(s)]$,when N is Binomial (m,p) then

$$F_N(s) = [q + ps]^m$$

where $q = 1 - p$
Thus

$$H(s) = [q + pG_X(s)]^m \tag{2.51}$$

When X is Bernoulli with parameter π,then

$$\begin{aligned} H(s) &= [q + pG_X(s)]^m \\ &= [q + p(1 - \pi + \pi s)]^m \\ &= [1 - p + p - p\pi + \pi ps]^m \\ &= [1 - p\pi + p\pi s]^m \end{aligned}$$

$$H(s) = [(1 - p\pi) + p\pi s]^m \tag{2.52}$$

which is the pgf of a Binomial distribution with parameters m and $p\pi$.

2.8.1.2 Compound Poisson Distribution

From $H(s) = F_N[G_X(s)]$,when N is Poisson (λ) then

$$F_N(s) = e^{-\lambda(1-s)}$$

Thus

$$H(s) = e^{-\lambda(1-G_X(s))} \tag{2.53}$$

When X is Bernoulli with parameter π,then

$$\begin{aligned} H(s) &= e^{-\lambda(1-G_X(s))} \\ &= e^{-\lambda(1-(1-\pi+\pi s))} \\ &= e^{-\lambda(1-1+\pi-\pi s)} \\ &= e^{-\lambda(\pi-\pi s)} \end{aligned}$$

Thus

$$H(s) = e^{-\lambda\pi(1-s)} \tag{2.54}$$

which is the pgf of a Poisson distribution with parameter $\pi\lambda$.

When X is Logarithmic with parameter p,then

$$G_X(s) = \frac{log(1-ps)}{log(1-p)}$$

Therefore

$$\begin{aligned} H(s) &= e^{-\lambda(1-G_X(s))} \\ &= e^{-\lambda[1-\frac{log(1-ps)}{log(1-p)}]} \\ &= e^{-\lambda[\frac{log(1-p)-log(1-ps)}{log(1-p)}]} \\ &= e^{-\lambda[\frac{log\frac{(1-p)}{(1-ps)}}{log(1-p)}]} \end{aligned}$$

i.e.

$$H(s) = e^{\frac{-\lambda}{log(1-p)}[log\frac{(1-p)}{(1-ps)}]} \tag{2.55}$$

Let $r = \frac{-\lambda}{log(1-p)}$
Therefore

$$H(s) = e^{r log \frac{(1-p)}{(1-ps)}}$$
$$= e^{log[\frac{(1-p)}{(1-ps)}]^r}$$

Thus

$$H(s) = [\frac{(1-p)}{(1-ps)}]^r \tag{2.56}$$

which is the pgf of the Negative Binomial distribution with parameters r and $(1-p)$.

2.8.1.3 Compound Geometric Distribution

From

$$H(s) = F_N[G_X(s)]$$

Type I:When N is the number of failures before the first success

$$F_N(s) = \frac{p}{1-qs}$$

where $q = 1 - p$
Thus

$$H(s) = \frac{p}{1-qG_X(s)} \tag{2.57}$$

When X is Bernoulli with parameter π,then

$$\begin{aligned}
H(s) &= \frac{p}{1-qG_X(s)} \\
&= \frac{p}{1-q(1-\pi+\pi s)} \\
&= \frac{p}{1-(1-p)(1-\pi+\pi s)} \\
&= \frac{p}{1-(1-\pi+\pi s-p+p\pi-p\pi s)} \\
&= \frac{p}{\pi-\pi s+p-p\pi+p\pi s)} \\
&= \frac{p}{p+\pi-p\pi-(1-p)\pi s)} \\
&= \frac{p}{p+(1-p)\pi-(1-p)\pi s)} \\
&= \frac{\frac{p}{p+(1-p)\pi}}{1-\frac{(1-p)\pi}{p+(1-p)\pi}s}
\end{aligned}$$

Let

$$\alpha = \frac{p}{p+(1-p)\pi}$$

Then

$$1-\alpha = 1-\frac{p}{p+(1-p)\pi} = \frac{(1-p)\pi}{p+(1-p)\pi}$$

Therefore

$$H(s) = \frac{\alpha}{1-(1-\alpha)s} \tag{2.58}$$

which is the pgf of a Geometric Distribution with parameter α.

Type II:When N is the total number of trials required to achieve the first s

$$F_N(s) = \frac{ps}{1-qs}$$

where $q = 1-p$
Thus

$$H(s) = \frac{ps}{1-qG_X(s)} \tag{2.59}$$

When X is Bernoulli with parameter π,then

$$
\begin{aligned}
H(s) &= \frac{ps}{1-qG_X(s)} \\
&= \frac{ps}{1-q(1-\pi+\pi s)} \\
&= \frac{ps}{1-(1-p)(1-\pi+\pi s)} \\
&= \frac{ps}{1-(1-\pi+\pi s-p+p\pi-p\pi s)} \\
&= \frac{ps}{\pi-\pi s+p-p\pi+p\pi s)} \\
&= \frac{ps}{p+\pi-p\pi-(1-p)\pi s)} \\
&= \frac{ps}{p+(1-p)\pi-(1-p)\pi s)} \\
&= \frac{\frac{p}{p+(1-p)\pi}s}{1-\frac{(1-p)\pi}{p+(1-p)\pi}s}
\end{aligned}
$$

Let

$$\alpha = \frac{p}{p+(1-p)\pi}$$

Then

$$1-\alpha = 1-\frac{p}{p+(1-p)\pi} = \frac{(1-p)\pi}{p+(1-p)\pi}$$

Therefore

$$H(s) = \frac{\alpha s}{1-(1-\alpha)s} \tag{2.60}$$

which is the pgf of a Geometric Distribution with parameter α.

2.8.1.4 Compound Negative Binomial Distribution

From

$$H(s) = F_N[G_X(s)]$$

Type I:When N is the number of failures before the r^{th} success

$$F_N(s) = [\frac{p}{1-qs}]^r$$

where $q = 1 - p$
Thus

$$H(s) \quad = \quad [\frac{p}{1 - qG_X(s)}]^r \tag{2.61}$$

When X is Bernoulli with parameter π,then

$$\begin{aligned} H(s) &= [\frac{p}{1 - qG_X(s)}]^r \\ &= [\frac{p}{1 - q(1 - \pi + \pi s)}]^r \\ &= [\frac{p}{1 - (1-p)(1 - \pi + \pi s)}]^r \\ &= [\frac{p}{1 - (1 - \pi + \pi s - p + p\pi - p\pi s)}]^r \\ &= [\frac{p}{\pi - \pi s + p - p\pi + p\pi s)}]^r \\ &= [\frac{p}{p + \pi - p\pi - (1-p)\pi s)}]^r \\ &= [\frac{p}{p + (1-p)\pi - (1-p)\pi s)}]^r \\ &= [\frac{\frac{p}{p+(1-p)\pi}}{1 - \frac{(1-p)\pi}{p+(1-p)\pi} s}]^r \end{aligned}$$

Let

$$\alpha = \frac{p}{p + (1-p)\pi}$$

Then

$$1 - \alpha = 1 - \frac{p}{p + (1-p)\pi} = \frac{(1-p)\pi}{p + (1-p)\pi}$$

Therefore

$$H(s) \quad = \quad [\frac{\alpha}{1 - (1-\alpha)s}]^r \tag{2.62}$$

which is the pgf of a Negative Binomial Distribution with parameters α and r.

Type II:When N is the total number of trials required to achieve r successe

$$F_N(s) = [\frac{ps}{1 - qs}]^r$$

where $q = 1 - p$
Thus

$$H(s) = [\frac{ps}{1 - qG_X(s)}]^r \tag{2.63}$$

When X is Bernoulli with parameter π,then

$$\begin{aligned}
H(s) &= [\frac{ps}{1 - qG_X(s)}]^r \\
&= [\frac{ps}{1 - q(1 - \pi + \pi s)}]^r \\
&= [\frac{ps}{1 - (1-p)(1 - \pi + \pi s)}]^r \\
&= [\frac{ps}{1 - (1 - \pi + \pi s - p + p\pi - p\pi s)}]^r \\
&= [\frac{ps}{\pi - \pi s + p - p\pi + p\pi s)}]^r \\
&= [\frac{ps}{p + \pi - p\pi - (1-p)\pi s)}]^r \\
&= [\frac{ps}{p + (1-p)\pi - (1-p)\pi s)}]^r \\
&= [\frac{\frac{p}{p+(1-p)\pi} s}{1 - \frac{(1-p)\pi}{p+(1-p)\pi} s}]^r
\end{aligned}$$

Let

$$\alpha = \frac{p}{p + (1-p)\pi}$$

Then

$$1 - \alpha = 1 - \frac{p}{p + (1-p)\pi} = \frac{(1-p)\pi}{p + (1-p)\pi}$$

Therefore

$$H(s) = [\frac{\alpha s}{1 - (1-\alpha)s}]^r \tag{2.64}$$

which is the pgf of a Negative Binomial Distribution with parameters α and r.

2.8.1.5 Compound Uniform Distribution

From $H(s) = F_N[G_X(s)]$ when N is Uniform(a),then

$$F_N(s) = \frac{s}{a}[\frac{1-s^a}{1-s}]$$

Thus

$$H(s) = \frac{G_X(s)}{a}[\frac{1-(G_X(s))^a}{1-G_X(s)}] \quad (2.65)$$

When X is Bernoulli with parameter π,then

$$\begin{aligned} H(s) &= \frac{(1-\pi+\pi s)}{a}[\frac{1-(1-\pi+\pi s)^a}{1-(1-\pi+\pi s)}] \\ &= \frac{(1-\pi+\pi s)}{a}[\frac{1-(1-\pi+\pi s)^a}{1-1+\pi-\pi s}] \\ &= \frac{(1-\pi+\pi s)}{a}[\frac{1-(1-\pi+\pi s)^a}{\pi-\pi s}] \\ &= \frac{(1-\pi+\pi s)}{a}[\frac{1-(1-\pi+\pi s)^a}{(1-s)\pi}] \end{aligned}$$

Let

$$\begin{aligned} z &= 1-\pi+\pi s \\ &= 1-\pi(1-s) \\ \Longrightarrow 1-z &= \pi(1-s) \end{aligned}$$

Thus

$$H(s) = \frac{z}{a}[\frac{1-z^a}{1-z}] \quad (2.66)$$

which is a pgf of a Uniform distribution.

2.8.1.6 Compound Logarithmic Distribution

From $H(s) = F_N[G_X(s)]$ when N is Logarithmic(q),then

$$F_N(s) = \frac{log(1-qs)}{log(1-q)}$$

Thus

$$H(s) = \frac{log(1 - qG_X(s))}{log(1 - q)} \tag{2.67}$$

When X is Bernoulli with parameter π,then

$$H(s) = \frac{log(1 - q(1 - \pi + \pi s))}{log(1 - q)}$$

$$= \frac{log(1 - q + q\pi - q\pi s)}{log(1 - q)} \tag{2.68}$$

which is a pgf but not of a Logarithmic distribution.

Chapter 3

MOMENTS BASED ON RECURSIVE RELATIONS

3.1 Introduction

Let

$$p_k = Prob(X = k) \;\; for \;\; k = 0, 1, 2, \ldots$$

This implies that

$$p_{k-1} = Prob(X = k - 1) \;\; for \;\; k = 1, 2, 3, \ldots$$

For zero-truncated distributions, p_k holds for $k = 1, 2, 3, \ldots$ and p_{k-1} is for $k = 2, 3, 4, \ldots$

The aim of this chapter is to determine the ratio p_k/p_{k-1} for each probability mass function under consideration.From this ratio we get the recursive relation by expressing p_k in terms of p_{k-1}.

Using this recursive relation we can obtain means and variances using the pgf techique,where possible,or by what we call Feller's method.

For pgf techique we define

$$G(s) = \sum_{k=0}^{\infty} p_k s^k = \sum_{k=1}^{\infty} p_{k-1} s^{k-1}$$

$$G'(s) = \frac{dG}{ds} = \sum_{k=1}^{\infty} k p_k s^{k-1} = \sum_{k=2}^{\infty} (k-1) p_{k-1} s^{k-2}$$

$$G''(s) = \frac{d^2G}{ds^2} = \sum_{k=2}^{\infty} k(k-1)p_{k-1}s^{k-2} = \sum_{k=3}^{\infty}(k-1)(k-2)p_{k-1}s^{k-3}$$

$$E(X) = G'(1)$$

$$Var(X) = G''(1) + G'(1) - [G'(1)]^2$$

For Feller's method

$$M_1 = E(X) = \sum_{k=1}^{\infty} kp_k$$

$$M_2 = E(X^2) = \sum_{k=1}^{\infty} k^2 p_k$$

$$Var(X) = M_2 - M_1^2$$

We are going to apply the above mentioned methods for the following standard discrete distributions along with some truncated distributions that have been identified in Chapter 2.

- Poisson Distribution
- Zero-truncated Poisson Distribution
- Geometric Distribution:Type I and II
- Zero-truncated Geometric Distribution
- Logarithmic Distribution
- Binomial Distribution:Type I and II
- Negative Binomial Distribution
- Zero-truncated Negative Binomial Distribution
- Hypergeometric Distribution

Also in Chapter 2,the following distinct compound discrete distributions have been derived.

- Beta-Geometric Distribution

- Beta-Binomial Distribution
- Beta-Negative Binomial Distribution

It is noted that a number of compound distributions have reverted back to standard discrete distributions.We have shown that
Binomial-Binomial is Binomial
Poisson-Binomial is Poisson
Exponential-Poisson is Geometric
Gamma-Poisson is Negative Binomial

3.2 Recursions and moments of Standard Discrete Distributions

3.2.1 Poisson Distribution

Given

$$p_k = \frac{e^{-\lambda}\lambda^k}{k!} \quad for \quad k = 0, 1, 2, \ldots$$

$$p_0 = e^{-\lambda}$$

then

$$p_{k-1} = \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \quad for \quad k = 1, 2, 3, \ldots$$

Therefore

$$\frac{p_k}{p_{k-1}} = \frac{\lambda}{k} \qquad (3.1a)$$

$$kp_k = \lambda p_{k-1} \quad for \quad k = 1, 2, 3, \ldots \qquad (3.1b)$$

To obtain M_1 sum (3.1b) over k to get

$$\sum_{k=1}^{\infty} kp_k = \lambda \sum_{k=1}^{\infty} p_{k-1}$$

$$\Longleftrightarrow M_1 = \lambda.1$$

Thus

$$E(X) = M_1 = \lambda$$

Multiply (3.1b) by k and sum the result over k

$$\sum_{k=1}^{\infty} k^2 p_k = \lambda \sum_{k=1}^{\infty} k p_{k-1}$$

i.e.,

$$\begin{aligned} M_2 &= \lambda \sum_{k=1}^{\infty} [(k-1)+1] p_{k-1} \\ &= \lambda \sum_{k=1}^{\infty} (k-1) p_{k-1} + \lambda \sum_{k=1}^{\infty} p_{k-1} \end{aligned}$$

Therefore

$$\begin{aligned} M_2 &= \lambda M_1 + \lambda \\ &= \lambda\lambda + \lambda \\ &= \lambda^2 + \lambda \end{aligned}$$

Thus

$$\begin{aligned} Var(X) &= M_2 - \{M_1\}^2 \\ &= \lambda^2 + \lambda - \lambda \\ &= \lambda \end{aligned}$$

For pgf, multiply (3.1b) by s^k and then sum the results over k

$$\sum_{k=1}^{\infty} k p_k s^k = \lambda \sum_{k=1}^{\infty} p_{k-1} s^k$$

$$\Longleftrightarrow s \sum_{k=1}^{\infty} k p_k s^{k-1} = \lambda s \sum_{k=1}^{\infty} p_{k-1} s^{k-1}$$

i.e.,

$$s \frac{dG}{ds} = \lambda s G(s)$$

where

$$G(s) = \sum_{k=0}^{\infty} p_k s^k = \sum_{k=1}^{\infty} p_{k-1} s^{k-1}$$

$$G'(s) = \frac{dG}{ds} = \sum_{k=1}^{\infty} kp_k s^{k-1} = \sum_{k=2}^{\infty} (k-1)p_{k-1}s^{k-2}$$

Therefore

$$\int \frac{dG}{G} = \lambda \int ds$$

$$\Longrightarrow logG(s) = \lambda s + k$$

$$G(s) = e^{\lambda s+k}$$

$$\Longrightarrow G(1) = e^{\lambda}e^{k}$$

$$\Longrightarrow 1 = e^{\lambda}e^{k}$$

$$\Longrightarrow e^{k} = C = e^{\lambda}$$

Hence

$$G(s) = e^{-\lambda(1-s)}$$

so

$$G'(s) = \lambda e^{-\lambda(1-s)}$$

and

$$G''(s) = \lambda^2 e^{-\lambda(1-s)}$$

Therefore

$$E(X) = G'(1) = \lambda$$

$$\begin{aligned} Var(X) &= G''(1) + G'(1) - [G'(1)]^2 \\ &= \lambda^2 + \lambda - \lambda^2 \\ &= \lambda \end{aligned}$$

3.2.2 Zero-Truncated Poisson Distribution

Given

$$p_k = (e^{-\lambda} - 1)^{-1}\frac{\lambda^k}{k} \qquad k = 1, 2, 3, \ldots$$

$$p_1 = \frac{\lambda}{e^{-\lambda} - 1}$$

then

$$p_{k-1} = (e^{-\lambda} - 1)^{-1}\frac{\lambda^{k-1}}{k-1} \qquad k = 2, 3, 4, \ldots$$

Therefore

$$\begin{aligned} \frac{p_k}{p_{k-1}} &= \frac{\lambda}{k} & (3.2a) \\ kp_k &= \lambda p_{k-1} \; for \; k = 2, 3, 4, \ldots & (3.2b) \end{aligned}$$

To obtain M_1 sum (3.2b) over k to get

$$\sum_{k=2}^{\infty} kp_k = \lambda \sum_{k=2}^{\infty} p_{k-1}$$

i.e.

$$M_1 - p_1 = \lambda$$

$$\Longrightarrow M_1 = \lambda + p_1$$

Thus

$$E(X) = M_1 = \lambda + \frac{\lambda}{(e^{-\lambda} - 1)}$$

Multiply (3.2b) by k and sum the results over k

$$\sum_{k=2}^{\infty} k^2 p_k = \lambda \sum_{k=2}^{\infty} kp_{k-1}$$

$$\Longleftrightarrow M_2 - p_1 = \lambda \sum_{k=2}^{\infty} [(k-1) + 1] p_{k-1}$$

$$= \lambda \sum_{k=2}^{\infty} (k-1) p_{k-1} + \lambda \sum_{k=1}^{\infty} p_{k-1}$$

i.e.

$$M_2 - p_1 = \lambda M_1 + \lambda$$

$$\Longrightarrow M_2 = \lambda M_1 + \lambda + p_1$$

Therefore

$$M_2 = \lambda^2 + \lambda + \frac{\lambda^2}{(e^{-\lambda} - 1)} + \frac{\lambda}{(e^{-\lambda} - 1)}$$

$$Var(X) = M_2 - \{M_1\}^2$$
$$= \lambda^2 + \lambda + \frac{\lambda^2}{(e^{-\lambda} - 1)} + \frac{\lambda}{(e^{-\lambda} - 1)} - [\lambda + \frac{\lambda}{(e^{-\lambda} - 1)}]^2$$
$$= \lambda + \frac{\lambda}{e^{-\lambda} - 1} - \frac{\lambda^2}{(e^{-\lambda} - 1)} - \frac{\lambda^2}{(e^{-\lambda} - 1)^2}$$

after simplification.
Therefore

$$Var(X) = \lambda + \frac{\lambda}{e^{-\lambda} - 1}(1 - \lambda - \frac{\lambda}{e^{-\lambda} - 1})$$

To obtain the pgf X, multiply (3.2b) by s^k then sum the results over k

$$\sum_{k=2}^{\infty} kp_k s^k = \lambda \sum_{k=2}^{\infty}(k-1)p_{k-1}s^k$$

$$\Longleftrightarrow s\sum_{k=2}^{\infty} kp_k s^{k-1} = \lambda s^2 \sum_{k=1}^{\infty}(k-1)p_{k-1}s^{k-2}$$

$$\Longrightarrow s[\frac{dG}{ds} - p_1] = \lambda s^2 \frac{dG}{ds})$$

i.e.

$$(1 - \lambda s)\frac{dG}{ds} = p_1 = \frac{\lambda}{e^{-\lambda} - 1}$$

Thus

$$\int dG = \frac{1}{e^{-\lambda} - 1}\int \frac{\lambda ds}{1 - \lambda s}$$

$$\Longrightarrow G(s) = -\frac{ln(1 - \lambda s)}{e^{-\lambda} - 1} + k$$

Putting $s = 1$ the pgf becomes

$$G(1) = -\frac{ln(1 - \lambda)}{e^{-\lambda} - 1} + k$$

$$\Longrightarrow k = 1 - \frac{ln(1 - \lambda)}{e^{-\lambda} - 1}$$

Therefore

$$G(s) = -\frac{ln(1 - \lambda)}{e^{-\lambda} - 1} + \frac{-ln(1 - \lambda)}{e^{-\lambda} - 1} + 1$$

so

$$G'(s) = \frac{dG}{ds} = \frac{\lambda}{(1-\lambda s)(e^{-\lambda}-1)}$$

and

$$G''(s) = \frac{d}{ds}G'(s) = \frac{\lambda^2}{(1-\lambda s)^2(e^{-\lambda}-1)}$$

Hence

$$E(X) = G'(1) = \frac{\lambda}{(1-\lambda)(e^{-\lambda}-1)}$$

$$G''(1) = \frac{\lambda^2}{(1-\lambda)^2(e^{-\lambda}-1)}$$

$$\begin{aligned} Variance(X) &= G''(1) + G'(1) - [G'(1)]^2 \\ &= \frac{\lambda^2}{(1-\lambda)^2(e^{-\lambda}-1)} + \frac{\lambda}{(1-\lambda)(e^{-\lambda}-1)} - [\frac{\lambda}{(1-\lambda)(e^{-\lambda}-1)}]^2 \\ &= \frac{\lambda e^{-\lambda} - \lambda - \lambda^2}{(1-\lambda)^2(e^{-\lambda}-1)^2} \end{aligned}$$

after simplification.
Therefore

$$Var(X) = \frac{\lambda e^{-\lambda} - \lambda - \lambda^2}{(1-\lambda)^2(e^{-\lambda}-1)^2}$$

3.2.3 Binomial Distribution

Given

$$p_k = \binom{n}{k} p^k q^{(n-k)} \qquad k = 0, 1, 2, \ldots, n \quad ; q = 1 - p$$

$$p_0 = q^n$$

then

$$p_{k-1} = \binom{n}{k-1} p^{k-1} q^{(n-(k-1))} \qquad k = 1, 2, \ldots, n$$

Therefore

$$\frac{p_k}{p_{k-1}} = \frac{n-k+1}{k}\frac{p}{q} \tag{3.3a}$$

$$kp_k = (n-k+1)\frac{p}{q}p_{k-1} \quad k = 1, 2, \ldots, n \tag{3.3b}$$

Using Feller's method
To obtain M_1 sum (3.3b) over k which results to

$$\sum_{k=1}^{n} kp_k = \sum_{k=1}^{n}(n-k+1)\frac{p}{q}p_{k-1}$$

i.e.;

$$\begin{aligned} M_1 &= \frac{p}{q}\sum_{k=1}^{n}[n-(k-1)]p_{k-1} \\ &= \frac{p}{q}[n\sum_{k=1}^{n}p_{k-1} - \sum_{k=1}^{n}(k-1)p_{k-1}] \\ &= \frac{p}{q}[n(1-p_n) - (M_1 - np_n)] \end{aligned}$$

Therefore

$$\begin{aligned} M_1 &= \frac{p}{q}[n - M_1] \\ \iff qM_1 &= p[n - M_1] \\ \implies M_1 &= np \end{aligned}$$

Thus

$$E(X) = M_1 = np$$

Multiply (3.3b) by k and sum the results over k

$$\sum_{k=1}^{n} k^2 p_k = \sum_{k=1}^{n} k(n-k+1)\frac{p}{q}p_{k-1}$$

$$\begin{aligned} \iff M_2 &= \frac{p}{q}\sum_{k=1}^{n}[(k-1)+1][n-(k-1)]p_{k-1} \\ &= \frac{p}{q}[n\sum_{k=1}^{n}(k-1)p_{k-1} - \sum_{k=1}^{n}(k-1)^2 p_{k-1} + n\sum_{k=1}^{n}p_{k-1} - \sum_{k=1}^{n}(k-1)p_{k-1}] \\ &= \frac{p}{q}[n(M_1 - np_n) - (M_2 - n^2 p_n) + n(1 - P_n) - (M_1 - np_n)] \end{aligned}$$

But $M_1 = np$ hence

$$M_2 = \frac{p}{1-p}[n^2 - M_2 + n - np]$$

$$\Longleftrightarrow qM_2 = p[n^2 - M_2 + n - np]$$

$$\Longrightarrow M_2 = n^2p^2 + np - np^2$$

Therefore

$$\begin{aligned} Var(X) &= M_2 - \{M_1\}^2 \\ &= n^2p^2 + np - np^2 - (np)^2 \\ &= np - np^2 \\ &= npq \end{aligned}$$

To obtain the pgf X
Multiply (3.3b) by s^k and sum the results over k

$$q\sum_{k=1}^{n} kp_k s^k = p\sum_{k=1}^{n}[n-(k-1)]p_{k-1}s^k$$

$$\Longleftrightarrow qs\sum_{k=1}^{\infty} kp_k s^{k-1} = p[n\sum_{k=1}^{\infty} p_{k-1}s^k - \sum_{k=1}^{\infty}(k-1)p_{k-1}s^k]$$

i.e.

$$qs\sum_{k=1}^{\infty} kp_k s^{k-1} = nps\sum_{k=1}^{\infty} p_{k-1}s^{k-1} - ps^2\sum_{k=1}^{\infty}(k-1)p_{k-1}s^{k-2}]$$

$$\Longrightarrow qs\frac{dG}{ds} = npsG(s) - ps^2\frac{dG}{ds}$$

$$\Longleftrightarrow (q+ps)\frac{dG}{ds} = npG(s)$$

$$\int\frac{dG}{G} = n\int\frac{pds}{(q+ps)}$$

$$\Longrightarrow lnG(s) = nln(q+ps) + lnk$$

Therefore

$$G(s) = k(q+ps)^n$$

Put $s = 1$

$$G(1) = k(q+p)^n$$
$$\Longrightarrow k = 1$$

Thus

$$G(s) = (q+ps)^n$$

Implying that

$$G'(s) = \frac{dG}{ds} = np(q+ps)^{n-1}$$

and

$$G''(s) = \frac{d}{ds}G'(s) = n(n-1)p^2(q+ps)^{n-2}$$

Hence

$$E(X) = G'(1) = np$$

and

$$G''(1) = n^2p^2 - np^2$$

Therefore

$$\begin{aligned} Var(X) &= G''(1) + G'(1) - [G'(1)]^2 \\ &= n^2p^2 - np^2 + np - n^2p^2 \\ &= np - np^2 \\ &= np(1-p) \\ &= npq \end{aligned}$$

3.2.4 Geometric Distribution

Type I: Let X be the number of failures before the first success

Given

$$p_k = q^k p \quad k = 0, 1, 2, \ldots \quad ; q = 1 - p$$
$$p_0 = p$$

then

$$p_{k-1} = Prob(X = k-1) = q^{k-1}p \quad k = 1, 2, 3, \ldots$$

$$\begin{aligned} \frac{p_k}{p_{k-1}} &= q & (3.4a) \\ p_k &= qp_{k-1} \quad k = 1, 2, \ldots & (3.4b) \end{aligned}$$

Using Feller's method
Multiply (3.4b) by k and sum the results over k

$$\sum_{k=1}^{\infty} kp_k = q\sum_{k=1}^{\infty} kp_{k-1}$$

$$\begin{aligned} M_1 &= q\sum_{k=1}^{\infty}[(k-1)+1]p_{k-1} \\ &= q[\sum_{k=1}^{\infty}(k-1)p_{k-1} + \sum_{k=1} p_{k-1}] \end{aligned}$$

$$\begin{aligned} M_1 &= qM_1 + q \\ \Longleftrightarrow M_1 - qM_1 &= q \\ \Longrightarrow M_1 &= \frac{q}{p} \end{aligned}$$

Thus

$$E(X) = M_1 = \frac{q}{p}$$

Multiply (3.4b) by k^2 and sum the results over k

$$\sum_{k=1}^{\infty} k^2 p_k = q\sum_{k=1}^{\infty} k^2 p_{k-1}$$

$$\begin{aligned} M_2 &= q\sum_{k=1}^{\infty}[(k-1)+1]^2 p_{k-1} \\ &= q\sum_{k=1}^{\infty}[(k-1)^2 + 2(k-1) + 1]p_{k-1} \\ &= q[\sum_{k=1}^{\infty}(k-1)^2 p_{k-1} + 2\sum_{k=1}^{\infty}(k-1)p_{k-1} + \sum_{k=1}^{\infty} p_{k-1}] \\ \Longrightarrow M_2 &= q[M_2 + 2M_1 + 1] \\ &= qM_2 + 2qM_1 + q \end{aligned}$$

But $M_1 = \frac{q}{p}$

$$\Longrightarrow M_2(1-q) = \frac{2q^2}{p} + q$$

$$M_2 = \frac{2q^2}{p^2} + \frac{q}{p}$$

$$\begin{aligned} Var(X) &= M_2 - \{M_1\}^2 \\ &= \frac{2q^2}{p^2} + \frac{q}{p} - (\frac{q}{p})^2 \\ &= \frac{q(q+p)}{p^2} \\ &= \frac{q}{p^2} \end{aligned}$$

since $q + p = 1$

To obtain the pgf of X

Multiply (3.4b) by s^k and sum the results over k

$$\sum_{k=1}^{\infty} p_k s^k = q \sum_{k=1}^{\infty} p_{k-1} s^k$$

$$\Longleftrightarrow G(s) - p_0 = qsG(s)$$

$$\Longrightarrow (1-qs)G(s) = p_0 = p$$

Thus

$$G(s) = \frac{p}{1-qs}$$

implying that

$$G'(s) = \frac{dG}{ds} = \frac{pq}{(1-qs)^2}$$

and

$$G''(s) = \frac{d}{ds}G'(s) = \frac{2pq^2}{(1-qs)^3}$$

Hence

$$\begin{aligned} E(X) = G'(1) &= \frac{pq}{(1-q)^2} \\ &= \frac{q}{p} \end{aligned}$$

since $1 - q = p$

$$G''(1) = \frac{2pq^2}{(1-q)^3} = \frac{2q^2}{p^2}$$

Therefore

$$\begin{aligned} Var(X) &= G''(1) + G'(1) - [G'(1)]^2 \\ &= \frac{2q^2}{p^2} + \frac{q}{p} - [\frac{q}{p}]^2 \\ &= \frac{q^2}{p^2} + \frac{q}{p} \\ &= \frac{q^2 + qp}{p^2} \\ &= \frac{q(q+p)}{p^2} \\ &= \frac{q}{p^2} \end{aligned}$$

Type II: Let X be the total number of trials required to achieve the first suc

Given

$$p_k = q^{k-1}p \quad k = 1, 2, \ldots \quad ; q = 1 - p$$

$$p_1 = p$$

then

$$p_{k-1} = q^{k-2}p \quad k = 2, 3, \ldots$$

Therefore

$$\frac{p_k}{p_{k-1}} = q \qquad (3.5a)$$

$$p_k = qp_{k-1} \quad k = 2, 3, \ldots \qquad (3.5b)$$

Using Feller's method

Multiply (3.5b) by k and sum the results over k

$$\sum_{k=2}^{\infty} kp_k = q\sum_{k=2}^{\infty} kp_{k-1}$$

$$[M_1 - p_1] = q\sum_{k=2}^{\infty}[(k-1)+1]p_{k-1}$$

$$= q[\sum_{k=2}^{\infty}(k-1)p_{k-1} + \sum_{k=2}^{\infty} p_{k-1}]$$

$$M_1 - p1 = qM_1 + q$$

but $p_1 = p$

$$\Longleftrightarrow M_1 - p = qM_1 + q$$

$$\Longrightarrow M_1 - qM_1 = q + p$$

$$pM_1 = 1$$

since $q + p = 1$
Thus

$$E(X) = M_1 = \frac{1}{p}$$

Multiplying (3.5b) by k^2 and summing the results over k

$$\sum_{k=2}^{\infty} k^2 p_k = q\sum_{k=2}^{\infty} k^2 p_{k-1}$$

$$[M_2 - p_1] = q\sum_{k=2}^{\infty}[(k-1)+1]^2 p_{k-1}$$

$$= q\sum_{k=2}^{\infty}[(k-1)^2 + 2(k-1) + 1]p_{k-1}$$

$$= q[\sum_{k=2}^{\infty}(k-1)^2 p_{k-1} + 2\sum_{k=2}^{\infty}(k-1)p_{k-1} + \sum_{k=2}^{\infty} p_{k-1}]$$

$$\Longrightarrow [M_2 - p_1] = q[M_2 + 2M_1 + 1]$$

But $p_1 = p$ and $M_1 = \frac{1}{p}$

$$M_2 - p = qM_2 + \frac{2q}{p} + q$$

$$\Longrightarrow M_2(1-q) = \frac{2q}{p} + q + p$$

But $1 - q = p$ and $q + p = 1$
Therefore

$$M_2 = \frac{2q}{p^2} + \frac{1}{p}$$

Hence

$$\begin{aligned} Var(X) &= M_2 - \{M_1\}^2 \\ &= \frac{2q}{p^2} + \frac{1}{p} - \{\frac{1}{p}\}^2 \\ &= \frac{q}{p^2} \end{aligned}$$

To obtain the pgf of X
Multiply (3.5b) by s^k and sum the results over k

$$\sum_{k=2}^{\infty} p_k s^k = q \sum_{k=2}^{\infty} p_{k-1} s^k$$

$$\Longleftrightarrow G(s) - p_1 s - p_0 = qs[G(s) - p_0]$$

But $p_1 = p$ and $p_0 = 0$

$$\Longrightarrow (1 - qs)G(s) = p_1 s = ps$$

Therefore

$$G(s) = \frac{ps}{1 - qs}$$

implying that

$$G'(s) = \frac{dG}{ds} = \frac{p}{(1 - qs)} + \frac{pqs}{(1 - qs)^2}$$

and

$$G''(s) = \frac{d}{ds}G'(s) = \frac{2qp}{(1 - qs)^2} + \frac{2pq^2 s}{(1 - qs)^3}$$

$$\begin{aligned} G'(1) &= \frac{p}{(1 - q)} + \frac{pq}{(1 - q)^2} \\ &= 1 + \frac{q}{p} \\ &= \frac{p + q}{p} \\ &= \frac{1}{p} \end{aligned}$$

Thus

$$E(X) = G'(1) = \frac{1}{p}$$

$$\begin{aligned} G''(1) &= \frac{2qp}{(1-q)^2} + \frac{2pq^2}{(1-q)^3} \\ &= \frac{2q}{p} + \frac{2q^2}{p^2} \end{aligned}$$

Hence

$$\begin{aligned} Var(X) &= G''(1) + G'(1) - [G'(1)]^2 \\ &= \frac{2q}{p} + \frac{2q^2}{p^2} + \frac{1}{p} - [\frac{1}{p}]^2 \\ &= \frac{2qp + 2q^2 + p - 1}{p^2} \\ &= \frac{2q(1-q) + 2q^2 + p - (q+p)}{p^2} \\ &= \frac{q}{p^2} \end{aligned}$$

3.2.5 Truncated Geometric Distribution

Given

$$p_k = q^{k-m}p \quad k = m, m+1, m+2, \ldots \quad ; q = 1-p$$

$$p_m = p$$

then

$$p_{k-1} = q^{(k-1)-m}p \quad k = m+1, m+2, \ldots$$

Therefore

$$\frac{p_k}{p_{k-1}} = q \qquad (3.6a)$$

$$p_k = qp_{k-1} \quad k = m+1, m+2, \ldots \qquad (3.6b)$$

Using Feller's method
Multiply (3.6b) by k and sum the results over k

$$\sum_{k=m+1}^{\infty} kp_k = q \sum_{k=m+1}^{\infty} kp_{k-1}$$

$$[M_1 - mp_m] = q\sum_{k=2}^{\infty}[(k-1)+1]p_{k-1}$$

$$= q[\sum_{k=2}^{\infty}(k-1)p_{k-1} + \sum_{k=2}^{\infty}p_{k-1}]$$

$$M_1 - mp_m = qM_1 + q$$

but $p_m = p$

$$\Longleftrightarrow M_1 - mp = qM_1 + q$$

$$\Longrightarrow M_1 - qM_1 = q + mp$$

$$\Longrightarrow pM_1 = q + mp$$

Thus

$$E(X) = M_1 = \frac{q+mp}{p} = m + \frac{q}{p}$$

Multiplying (3.6b) by k^2 and summing the results over k

$$\sum_{k=m+1}^{\infty} k^2 p_k = q \sum_{k=m+1}^{\infty} k^2 p_{k-1}$$

$$[M_2 - m^2 p_m] = q \sum_{k=m+1}^{\infty} [(k-1)+1]^2 p_{k-1}$$

$$= q \sum_{k=m+1}^{\infty} [(k-1)^2 + 2(k-1) + 1] p_{k-1}$$

$$= q[\sum_{k=m+1}^{\infty} (k-1)^2 p_{k-1} + 2\sum_{k=2}^{\infty}(k-1)p_{k-1} + \sum_{k=2}^{\infty} p_{k-1}]$$

$$\Longrightarrow [M_2 - m^2 p_m] = q[M_2 + 2M_1 + 1]$$

But $p_m = p$ and $M_1 = m + \frac{q}{p}$

$$M_2 - m^2 p = qM_2 + 2q(m + \frac{q}{p}) + q$$

$$\Longrightarrow M_2(1-q) = 2qm + \frac{2q^2}{p} + q + m^2 p$$

But $1 - q = p$
Therefore

$$M_2 = \frac{2qm}{p} + \frac{2q^2}{p^2} + \frac{q}{p} + m^2$$

Hence

$$\begin{aligned}
Var(X) &= M_2 - \{M_1\}^2 \\
&= \frac{2qm}{p} + \frac{2q^2}{p^2} + \frac{q}{p} + m^2 - \{m + \frac{q}{p}\}^2 \\
&= \frac{2qm}{p} + \frac{2q^2}{p^2} + \frac{q}{p} + m^2 - m^2 - 2m\frac{q}{p} - \frac{q^2}{p^2} \\
&= \frac{q^2}{p^2} + \frac{q}{p} \\
&= \frac{q^2 + qp}{p^2} \\
&= \frac{q(q+p)}{p^2} \\
&= \frac{q}{p^2}
\end{aligned}$$

To obtain the pgf of X
Multiply (3.6b) by s^k and sum the results over k

$$\sum_{k=m+1}^{\infty} p_k s^k = q \sum_{k=m+1}^{\infty} p_{k-1} s^k$$

$$\iff G(s) - p_m s^m = qsG(s)$$

But $p_m = p$

$$\implies (1 - qs)G(s) = p_m s^m = ps^m$$

Therefore

$$G(s) = \frac{ps^m}{1 - qs}$$

implying that

$$G'(s) = \frac{dG}{ds} = \frac{mps^{m-1}}{(1 - qs)} + \frac{pqs^m}{(1 - qs)^2}$$

and

$$G''(s) = \frac{d}{ds}G'(s) = \frac{m(m-1)ps^{m-2}}{(1-qs)} + \frac{2pqms^{m-1}}{(1-qs)^2} + \frac{2pq^2s^m}{(1-qs)^3}$$

Hence

$$\begin{aligned} G'(1) &= \frac{mp}{(1-q)} + \frac{pq}{(1-q)^2} \\ &= \frac{mp}{p} + \frac{qp}{p^2} \\ &= m + \frac{q}{p} \end{aligned}$$

Thus

$$E(X) = G'(1) = m + \frac{q}{p}$$

$$\begin{aligned} G''(1) &= \frac{m(m-1)p}{(1-q)} + \frac{2pqm}{(1-q)^2} + \frac{2pq^2}{(1-q)^3} \\ &= \frac{m(m-1)p}{p} + \frac{2pqm}{p^2} + \frac{2pq^2}{p^3} \\ &= m(m-1) + \frac{2qm}{p} + \frac{2pq^2}{p^2} \end{aligned}$$

Hence

$$\begin{aligned} Var(X) &= G''(1) + G'(1) - [G'(1)]^2 \\ &= m(m-1) + \frac{2qm}{p} + \frac{2pq^2}{p^3} - [m + \frac{q}{p}]^2 \\ &= m^2 - m + \frac{2qm}{p} + \frac{2q^2}{p^2} + m + \frac{q}{p} - m^2 - 2m\frac{q}{p} - \frac{q^2}{p^2} \\ &= \frac{q^2}{p^2} + \frac{q}{p} \\ &= \frac{q^2 + qp}{p^2} \\ &= \frac{q(q+p)}{p^2} \\ &= \frac{q}{p^2} \end{aligned}$$

3.2.6 Negative Binomial Distribution

Type I: Let X be the number of failures before the r^{th} success

Given

$$p_k = \binom{k+r-1}{r-1} q^k p^r \quad k = 0, 1, 2, \ldots \quad ; q = 1 - p$$

$$p_0 = p^r$$

then

$$p_{k-1} = \binom{(k-1)+r-1}{r-1} q^{k-1} p^r \quad k = 1, 2, 3, \ldots$$

$$\frac{p_k}{p_{k-1}} = q\frac{k+r-1}{k} \qquad (3.7a)$$

$$kp_k = q(k+r-1)p_{k-1} \quad k = 1, 2, 3, \ldots \qquad (3.7b)$$

Using Feller's method
To obtaim M_1 sum (3.7b) over k i.e.;

$$\sum_{k=1}^{\infty} kp_k = q\sum_{k=1}^{\infty}(k+r-1)p_{k-1}$$

$$M_1 = q\sum_{k=1}^{\infty}[(k-1)+r]p_{k-1}$$

$$= q\sum_{k=1}^{\infty}(k-1)p_{k-1} + qr\sum_{k=1}^{\infty}p_{k-1}$$

$$M_1 = qM_1 + qr$$

$$M_1(1-q) = qr$$

Thus

$$E(X) = M_1 = \frac{qr}{p}$$

Multiply (3.7b) by k sum the results over k

$$\sum_{k=1}^{\infty} k^2 p_k = q\sum_{k=1}^{\infty} k(k+r-1)p_{k-1}$$

$$
\begin{aligned}
M_2 &= q\sum_{k=1}^{\infty}[(k-1)+1][(k-1)+r]p_{k-1} \\
&= q[\sum_{k=1}^{\infty}(k-1)^2 p_{k-1} + (r+1)\sum_{k=1}^{\infty}(k-1)p_{k-1} + r\sum_{k=1}^{\infty}p_{k-1}] \\
&= qM_2 + q(r+1)M_1 + qr
\end{aligned}
$$

But $M_1 = \frac{qr}{p}$

$$
\begin{aligned}
\Longrightarrow M_2(1-q) &= (qr)^2 + q^2r + qr \\
&= \frac{(qr)^2 + q^2r + qr}{p}
\end{aligned}
$$

$$
\begin{aligned}
Var(X) &= M_2 - \{M_1\}^2 \\
&= \frac{(qr)^2 + q^2r + qr}{p} - \{\frac{qr}{p}\}^2 \\
&= \frac{qr}{p^2}
\end{aligned}
$$

To obtain the pgf
Multiplying (3.7b) by s^k and sum the results over k

$$
\begin{aligned}
\sum_{k=1}^{\infty} kp_k s^k &= q\sum_{k=1}^{\infty}[r+(k-1)]p_{k-1}s^k \\
&= qr\sum_{k=1}^{\infty}p_{k-1}s^k + q\sum_{k=1}^{\infty}(k-1)p_{k-1}s^k
\end{aligned}
$$

$$
\begin{aligned}
\Longleftrightarrow s\sum_{k=1}^{\infty}kp_k s^{k-1} &= qrs\sum_{k=1}^{\infty}p_{k-1}s^{k-1} + qs^2\sum_{k=1}^{\infty}(k-1)p_{k-1}s^{k-2} \\
\Longrightarrow qs\frac{dG}{ds} &= qrsG(s) + qs^2\frac{dG}{ds}
\end{aligned}
$$

i.e.

$$
(1-qs)\frac{dG}{ds} = qrG(s)
$$

Thus

$$\int \frac{dG}{G} = \int r\frac{qds}{(1-qs)}$$
$$\Longrightarrow lnG(s) = -rln(1-qs) + lnk$$

Therefore

$$G(s) = k(1-qs)^{-r}$$

Put $s = 1$

$$G(1) = k(1-q)^{-r} = kp^{-r}$$
$$\Longrightarrow k = p^r$$

since $G(1) = 1$ Thus

$$G(s) = p^r(1-qs)^{-r} = [\frac{p}{1-qs}]^r$$

implying that

$$G'(s) = \frac{qrp^r}{(1-qs)^{r+1}}$$

and

$$G''(s) = \frac{q^2r(r+1)p^r}{(1-qs)^{r+2}}$$
$$= \frac{q^2r^2p^r + q^2rp^r}{(1-qs)^{r+2}}$$

Thus

$$E(X) = G'(1) = \frac{qr}{p}$$

$$G''(1) = q^2r(r+1)p^r(1-q)^{-r-2}$$
$$= \frac{q^2r^2 + q^2r}{p^2}$$

Therefore

$$Var(X) = G''(1) + G'(1) - [G'(1)]^2$$
$$= \frac{q^2r^2 + q^2r}{p^2} + \frac{qr}{p} - \frac{q^2r^2}{p^2}$$
$$= \frac{qr}{p^2}$$

Type II: Let X be the number of trials required to achieve the r^{th} success

Given

$$p_k = \binom{k-1}{r-1} q^{k-r} p^r \quad k = r, r+1, r+2, \ldots \; ; q = 1-p$$

$$p_r = p^r$$

then

$$p_{k-1} = \binom{(k-1)-1}{r-1} q^{(k-1)-r} p^r \quad k = r+1, r+2, \ldots$$

$$\frac{p_k}{p_{k-1}} = q\frac{k-1}{k-r} \qquad (3.8a)$$

$$(k-r)p_k = q(k-1)p_{k-1} \quad k = r+1, r+2, \ldots \qquad (3.8b)$$

Using Feller's method

Multiply (3.8b) by k and sum the results over k

$$\sum_{k=r+1}^{\infty} kp_k = q \sum_{k=r+1}^{\infty} (k-1)p_{k-1} + r \sum_{k=r+1}^{\infty} p_k$$

$$\iff M_1 - rp_r = qM_1 + r(1-p_r)$$

i.e.

$$M_1(1-q) = r$$

But $1 - q = p$

Therefore

$$E(X) = M_1 = \frac{r}{p}$$

Multiplying (3.8b) by k^2 and sum the results over k

$$\sum_{k=r+1}^{\infty} k^2 p_k = q \sum_{k=r+1}^{\infty} [(k-1)+1](k-1)p_{k-1} + r \sum_{k=r+1}^{\infty} kp_k$$

$$M_2 - r^2 p_r = q[\sum_{k=r+1}^{\infty} (k-1)^2 p_{k-1} + \sum_{k=r+1}^{\infty} (k-1)p_{k-1}] + r(M_1 - rp_r)$$

$$= qM_2 + qM_1 + rM_1 - r^2 p_r$$

But

$$M_1 = \frac{r}{p}$$

Thus

$$M_2(1-q) = \frac{qr + r^2}{p}$$

$$M_2 = \frac{qr + r^2}{p^2}$$

since $1 - q = p$

$$\begin{aligned} Var(X) &= M_2 - [M_1]^2 \\ &= \frac{qr + r^2}{p^2} - [\frac{r}{p}]^2 \\ &= \frac{qr}{p^2} \end{aligned}$$

To obtain the pgf
Multiply(3.8b) by s^k and sum the results over k

$$\sum_{k=r+1}^{\infty} kp_k s^k - r \sum_{k=r+1}^{\infty} p_k s^k = q \sum_{k=r+1}^{\infty} (k-1)p_{k-1} s^k$$

$$\Longleftrightarrow s \sum_{k=r+1}^{\infty} kp_k s^{k-1} - r \sum_{k=r+1}^{\infty} p_k s^k = qs^2 \sum_{k=r+1}^{\infty} (k-1)p_{k-1} s^{k-2}$$

$$\Longrightarrow s[\frac{dG}{ds} - rp^r s^{r-1}] - r[G(s) - p^r s^r] = qs^2 \frac{dG}{ds}$$

i.e.

$$s(1-qs)\frac{dG}{ds} = rG(s)$$

$$\int \frac{dG}{ds} = r \int \frac{ds}{s(1-qs)}$$

Applying the partial fractions technique,

$$\int \frac{dG}{ds} = r[\int \frac{ds}{s} + \int \frac{qds}{(1-qs)}]$$

$$\Longrightarrow lnG(s) = r[lns - ln(1-qs)] + lnk$$

$$\Longleftrightarrow lnG(s) = rln[\frac{s}{1-qs}] + lnk$$

Thus

$$G(s) = k[\frac{s}{1-qs}]^r$$

Put $s = 1$

$$G(1) = k[\frac{1}{1-q}]^r$$

$$\Longrightarrow k = (1-q)^r = p^r$$

since $G(1) = 1$ and $p = 1 - q$
Therefore

$$G(s) = p^r[\frac{s}{1-qs}]^r$$
$$= [\frac{ps}{1-qs}]^r$$

implying that

$$G'(s) = \frac{dG}{ds} = \frac{rp^r s^{r-1}}{(1-qs)^r} + \frac{qrp^r s^r}{(1-qs)^{r+1}}$$

and

$$G''(s) = \frac{r(r-1)p^r s^{r-2}}{(1-qs)^r} + \frac{2qr^2p^r s^{r-1}}{(1-qs)^{r+1}} + \frac{q^2r(r+1)p^r s^r}{(1-qs)^{r+2}}$$

Put $s = 1$

$$G'(1) = \frac{rp^r}{(1-q)^r} + \frac{qrp^r}{(1-q)^{r+1}}$$
$$= \frac{rp^r}{p^r} + \frac{qrp^r}{p^{r+1}}$$
$$= \frac{r}{p}$$

Therefore

$$E(X) = G'(1) = \frac{r}{p}$$

$$G''(1) = \frac{r(r-1)p^r}{(1-q)^r} + \frac{2qr^2p^r}{(1-q)^{r+1}} + \frac{q^2r(r+1)p^r}{(1-q)^{r+2}}$$
$$= \frac{r(r-1)p^r}{p^r} + \frac{2qr^2p^r}{p^{r+1}} + \frac{q^2r(r+1)p^r}{p^{r+2}}$$
$$= r(r-1) + \frac{2qr^2}{p} + \frac{q^2r(r+1)}{p^2}$$

$$
\begin{aligned}
Var(X) &= G''(1) + G'(1) - [G'(1)]^2 \\
&= r(r-1) + \frac{2qr^2}{p} + \frac{q^2 r(r+1)}{p^2} + \frac{r}{p} - [\frac{r}{p}]^2 \\
&= \frac{r^2p^2 - rp^2 + 2qr^2p + q^2r^2 + q^2r + rp - r^2}{p^2} \\
&= \frac{qr}{p^2}
\end{aligned}
$$

3.2.7 Truncated Negative Binomial Distribution

Given

$$p_k = \frac{p^r}{1-p^r}\binom{k+r-1}{r-1} q^k p^r \quad k = 1, 2, 3, \ldots; \quad q = 1-p$$

$$p_1 = \frac{qrp^r}{1-p^r}$$

then

$$p_{k-1} = \frac{p^r}{1-p^r}\binom{(k-1)+r-1}{r-1} q^{k-1} p^r \quad k = 1, 2, 3, \ldots$$

$$
\begin{aligned}
\frac{p_k}{p_{k-1}} &= q\frac{k+r-1}{k} && (3.9a) \\
kp_k &= q(k+r-1)p_{k-1} \quad k = 2, 3, 4, \ldots && (3.9b)
\end{aligned}
$$

Using Feller's method
Summing (3.9b) over k results to

$$\sum_{k=2}^{\infty} kp_k = q\sum_{k=2}^{\infty}(k+r-1)p_{k-1}$$

$$
\begin{aligned}
M_1 - p_1 &= q\sum_{k=2}^{\infty}[(k-1)+r]p_{k-1} \\
&= qr\sum_{k=2}^{\infty} p_{k-1} + q\sum_{k=2}^{\infty}(k-1)p_{k-1}
\end{aligned}
$$

i.e.

$$M_1 - p_1 = qr + qM_1$$

$$\Longrightarrow (1-q)M_1 = qr + p_1$$
$$\Longleftrightarrow M_1 = \frac{qr + p_1}{p}$$

But $p_1 = \frac{qrp^r}{1-p^r}$
Therefore

$$E(X) = M_1 = \frac{qr}{p} + \frac{qrp^r}{1-p^r}$$
$$= \frac{qr}{p(1-p^r)}$$

Multiply (3.9b) by k and sum the results over k

$$\sum_{k=2}^{\infty} k^2 p_k = q\sum_{k=2}^{\infty} k(k+r-1)p_{k-1}$$

$$M_2 - p_1 = q\sum_{k=2}^{\infty}[(k-1)+1][(k-1)+r]p_{k-1}$$
$$= q[\sum_{k=2}^{\infty}(k-1)^2 p_{k-1} + (r+1)\sum_{k=2}^{\infty}(k-1)p_{k-1} + r\sum_{k=2}^{\infty} p_{k-1}]$$
$$= qM_2 + (qr+q)M_1 + qr$$

$$\Longrightarrow M_2(1-q) = (qr)^2 + q^2 r + qr + p_1$$
$$= (qr)^2 + q^2 r + qr + \frac{qrp^r}{1-p^r}$$

since $p_1 = \frac{qrp^r}{1-p^r}$ and $p = 1-q$

$$\Longleftrightarrow M_2 = \frac{(qr)^2 + q^2 r + qr}{p} + \frac{qrp^r}{p(1-p^r)}$$

$$Var(X) = M_2 - \{M_1\}^2$$
$$= \frac{(qr)^2 + q^2 r + qr}{p} + \frac{qrp^r}{p(1-p^r)} - \{\frac{qr}{p(1-p^r)}\}^2$$
$$= \frac{qr}{p^2(1-p^r)} - \frac{q^2 r^2 p^r}{p^2(1-p^r)^2}$$

To obtain the pgf,
Multiply (3.9b) by s^k and sum the results over k

$$\sum_{k=2}^{\infty} kp_k s^k = q\sum_{k=2}^{\infty}(k+r-1)p_{k-1}s^k$$
$$= q\sum_{k=2}^{\infty}[(k-1)+r]p_{k-1}s^k$$
$$= q\sum_{k=2}^{\infty}(k-1)p_{k-1}s^k + qr\sum_{k=2}^{\infty}p_{k-1}s^k$$

$$\Longleftrightarrow s\sum_{k=2}^{\infty} kp_k s^{k-1} = qs^2\sum_{k=2}^{\infty}(k-1)p_{k-1}s^{k-2} + qrs\sum_{k=2}^{\infty}p_{k-1}s^{k-1}$$

$$\Longrightarrow s[\frac{dG}{ds} - p_1] = qs^2\frac{dG}{ds} + qrs[G(s) - p_0]$$

But $p_0 = 0$ and $p_1 = \frac{qrp^r}{1-p^r}$
Therefore

$$s\frac{dG}{ds} - \frac{qrp^r}{1-p^r}s = qs^2\frac{dG}{ds} + qrsG(s)$$

i.e.;

$$s(1-qs)\frac{dG}{ds} = qrsG(s) + \frac{qrp^r}{1-p^r}s$$

which cannot be solved explicitly.

3.2.8 Logarithmic Distribution

Given

$$p_k = \frac{p^k}{-klogq} \quad k = 1, 2, \ldots; \quad q = 1-p$$

$$p_1 = \frac{-p}{logq}$$

then

$$p_{k-1} = \frac{p^{k-1}}{-(k-1)logq} \quad k = 2, 3, \ldots$$

$$\frac{p_k}{p_{k-1}} = p(\frac{k-1}{k}) \quad (3.10a)$$

$$\Longrightarrow kp_k = p(k-1)p_{k-1} \quad k = 2, 3, \ldots \quad (3.10b)$$

Using Feller's method
Summing (3.10b) over k results to

$$\sum_{k=2}^{\infty} kp_k = p\sum_{k=2}^{\infty}(k-1)p_{k-1}$$

$$\Longleftrightarrow M_1 - p_1 = pM_1$$

But

$$p_1 = \frac{-p}{logq}$$

Therefore

$$M_1(1-p) = \frac{-p}{logq}$$

$$E(X) = M_1 = \frac{-p}{qlogq}$$

since $1 - p = q$

Multiply (3.10b) by k and sum the results over k

$$\sum_{k=2}^{\infty} k^2 p_k = p\sum_{k=2}^{\infty} k(k-1)p_{k-1}$$

$$\begin{aligned} M_2 - p_1 &= p\sum_{k=2}^{\infty}[(k-1)+1](k-1)p_{k-1} \\ &= p[\sum_{k=2}^{\infty}(k-1)^2 p_{k-1} + \sum_{k=2}^{\infty}(k-1)p_{k-1}] \\ &= p[M_2 + M_1] \end{aligned}$$

$$\Longrightarrow M_2 - p_1 = pM_2 + pM_1$$

But $M_1 = \frac{-p}{qlogq}$ and $P_1 = \frac{-p}{logq}$
Therefore

$$M_2(1-p) = \frac{-p}{logq} + \frac{-p^2}{qlogq}$$

But $1-p=q$

$$M_2 = \frac{-p}{qlogq} + \frac{-p^2}{q^2logq}$$

$$\begin{aligned} Var(X) &= M_2 - [M_1]^2 \\ &= \frac{-p}{qlogq} + \frac{-p^2}{q^2logq} - (\frac{-p}{qlogq})^2 \\ &= \frac{-plogq - p^2}{q^2(logq)^2} \end{aligned}$$

To obtain the pgf
Multiply (3.10b) by s^k and sum the results over k

$$\sum_{k=2}^{\infty} kp_k s^k = p\sum_{k=2}^{\infty}(k-1)p_{k-1}s^k$$

$$s\sum_{k=2}^{\infty} kp_k s^{k-1} = ps^2\sum_{k=2}^{\infty}(k-1)p_{k-1}s^{k-2}$$

$$\Longrightarrow s[\frac{dG}{ds} - p_1] = ps^2\frac{dG}{ds}$$

i.e.

$$(1-ps)\frac{dG}{ds} = \frac{-p}{logq}$$

since

$$p_1 = \frac{-p}{logq}$$

Therefore

$$\int dG = \frac{1}{logq}\int \frac{-pds}{1-ps}$$

Thus

$$G(s) = \frac{log(1-ps)}{logq}$$

implying that

$$G'(s) = \frac{-p}{(1-ps)logq}$$

and

$$G''(s) = \frac{d}{ds}G'(s) = \frac{-p^2}{(1-ps)^2logq}$$

$$E(X) = G'(1) = \frac{-p}{(1-p)logq}$$
$$= \frac{-p}{qlogq}$$

$$G''(1) = \frac{-p^2}{(1-p)^2 logq}$$
$$= \frac{-p^2}{q^2 logq}$$

Therefore

$$Var(X) = G''(1) + G'(1) - [G'(1)]^2$$
$$= \frac{-p^2}{q^2 logq} + \frac{-p}{(1-p)logq} - \frac{-p^2}{q^2(logq)^2}$$
$$= \frac{-plogq - p^2}{q^2(logq)^2}$$

3.2.9 Hypergeometric Distribution

Given

$$p_k = Prob(X = k) = \frac{\binom{d}{k}\binom{N-d}{n-k}}{\binom{N}{n}} \quad k = 0, 1, 2, \ldots, d$$
$$p_0 = \frac{\binom{N-d}{n}}{\binom{N}{n}}$$

then

$$p_{k-1} = Prob(X = k-1) = \frac{\binom{d}{k-1}\binom{N-d}{n-(k-1)}}{\binom{N}{n}} \quad k = 1, 2, \ldots, d$$

$$\frac{p_k}{p_{k-1}} = \frac{(d-k+1)(n-k+1)}{k(N-d-n+k)} \quad (3.11a)$$
$$k(N-d-n+k)p_k = (d-k+1)(n-k+1)p_{k-1} \quad k = 1, 2, \ldots, d \quad (3.11b)$$

Using Feller's method,
Summing (3.11b) over k results to

$$\sum_{k=1}^{d} k(N-d-n+k)p_k = \sum_{k=1}^{d}(d-k+1)(n-k+1)p_{k-1}$$

$$(N-d-n)\sum_{k=1}^{d} kp_k + \sum_{k=1}^{d} k^2 p_k = \sum_{k=1}^{d}[d-(k-1)][n-(k-1)]p_{k-1}$$

$$\begin{aligned}\iff (N-d-n)M_1 + M_2 &= dn\sum_{k=1}^{d} p_{k-1} - (d+n)\sum_{k=1}^{d}(k-1)p_{k-1} + \sum_{k=1}^{d}(k-1)^2 p_{k-1} \\ &= dn[1-p_d] - (d+n)[M_1 - dp_d] + [M_2 - d^2 p_d] \\ &\implies NM_1 = dn\end{aligned}$$

Therefore

$$E(X) = M_1 = \frac{dn}{N}$$

Multiply (3.11b) by k and sum the results over k

$$\sum_{k=1}^{d} k^2(N-d-n+k)p_k = \sum_{k=1}^{d} k(d-k+1)(n-k+1)p_{k-1}$$

$$(N-d-n)\sum_{k=1}^{d} k^2 p_k + \sum_{k=1}^{d} k^3 p_k = \sum_{k=1}^{d}[(k-1)+1][d-(k-1)][n-(k-1)]p_{k-1}$$

This is equivalent to

$$\begin{aligned}(N-d-n)M_2 + M_3 &= (dn-d-n)\sum_{k=1}^{d}(k-1)p_{k-1} - (d+n-1)\sum_{k=1}^{d}(k-1)^2 p_{k-1} \\ &\quad + \sum_{k=1}^{d}(k-1)^3 p_{k-1} + dn\sum_{k=1}^{d} p_{k-1}\end{aligned}$$

i.e.

$$\begin{aligned}(N-d-n)M_2 + M_3 &= dn[M_1 - dp_d] - (d+n)[M_2 - d^2 p_d] + [M_3 - d^3 p_d] + dn[1-p_d] \\ &\quad -(d+n)[M_1 - dp_d] + [M_2 - d^2 p_d]\end{aligned}$$

$$\Longrightarrow M_2(N-1) = dn + (dn - d - n)M_1$$

But $M_1 = \frac{dn}{N}$
Hence

$$M_2 = \frac{Ndn + dn(dn - d - n)}{N(N-1)}$$

Therefore

$$\begin{aligned} Var(X) &= M_2 - [M_1]^2 \\ &= \frac{Ndn + dn(dn - d - n)}{N(N-1)} - [\frac{dn}{N}]^2 \\ &= \frac{d^2n^2 + dnN^2 - d^2nN - dn^2N}{N^2(N-1)} \\ &= \frac{dn(N-n)(N-d)}{N^2(N-1)} \end{aligned}$$

To obtain the pgf
Multiply (3.11b) by s^k and sum the results over k

$$\sum_{k=1}^{d} k(N - d - n + k)p_k s^k = \sum_{k=1}^{d} (d - k + 1)(n - k + 1)p_{k-1}s^k$$

$$\begin{aligned} (N-d-n)\sum_{k=1}^{d} kp_k s^k + \sum_{k=1}^{d} k^2 p_k s^k &= \sum_{k=1}^{d} [d-(k-1)][n-(k-1)]p_{k-1}s^k \\ &= nd\sum_{k=1}^{d} p_{k-1}s^k - (n+d)\sum_{k=1}^{d}(k-1)p_{k-1}s^k \\ &\quad + \sum_{k=1}^{d}(k-1)^2 p_{k-1}s^k \end{aligned}$$

which is equivalent to

$$\begin{aligned} (N-d-n)s\sum_{k=1}^{d} kp_k s^{k-1} + \sum_{k=1}^{d} k^2 p_k s^k &= nds\sum_{k=1}^{d} p_{k-1}s^{k-1} - (n+d)s^2\sum_{k=1}^{d}(k-1)p_{k-1}s^{k-2} \\ &\quad + \sum_{k=1}^{d}(k-1)^2 p_{k-1}s^k \end{aligned}$$

i.e.;

$$(N-d-n)s\frac{dG}{ds}+s^2\frac{d^2G}{ds^2}+s\frac{dG}{ds}=nds[G(s)-p_ds^d]-(n+d)s^2[\frac{dG}{ds}-dp_ds^{d-1}]+s^3\frac{d^2G}{ds^2}+s^2\frac{dG}{ds}$$

$$\implies [(N-d-n)s+(n+d-1)s^2]\frac{dG}{ds}+s^2(1-s)\frac{d^2G}{ds^2}=ndsG(s)+d^2p_ds^{d+1}$$

which cannot be solved explicitly.

3.3 Recursions and moments for Compound Distributions

3.3.1 Beta - Geometric Distribution

Type I: When X is the number of failures before the first success

Given

$$p_k=\frac{\Gamma(\alpha+1)\Gamma(\alpha+\beta)\Gamma(\beta+k)}{\Gamma\beta\Gamma(\alpha+\beta+k+1)}\qquad k=0,1,2,\dots$$

$$p_0=\frac{\Gamma(\alpha+1)}{(\alpha+\beta)}$$

then

$$p_{k-1}=\frac{\Gamma(\alpha+1)\Gamma(\alpha+\beta)\Gamma(\beta+(k-1))}{\Gamma\beta\Gamma(\alpha+\beta+(k-1)+1)}\qquad k=0,1,2,\dots$$

$$\frac{p_k}{p_{k-1}}=\frac{(\beta+k-1)}{(\alpha+\beta+k)}\qquad(3.12a)$$

$$\implies(\alpha+\beta+k)p_k=(\beta+k-1)p_{k-1}\qquad k=1,2,\dots\qquad(3.12b)$$

Using Feller's method

Multiplying (3.12b) by k and sum the results over k

$$\sum_{k=1}^{\infty}k(\alpha+\beta+k)p_k=\sum_{k=1}^{\infty}k(\beta+k-1)p_{k-1}$$

$$(\alpha+\beta)\sum_{k=1}^{\infty}kp_k+\sum_{k=1}^{\infty}k^2p_k=\sum_{k=1}^{\infty}[(k-1)+1][\beta+(k-1)]p_{k-1}$$

$$(\alpha+\beta)M_1+M_2=\beta\sum_{k=1}^{\infty}(k-1)p_{k-1}+\sum_{k=1}^{\infty}(k-1)^2p_{k-1}+\beta\sum_{k=1}^{\infty}p_{k-1}+\sum_{k=1}^{\infty}(k-1)p_{k-1}$$

$$\Longrightarrow (\alpha+\beta)M_1+M_2=\beta M_1+M_2+\beta+M_1$$

$$E(X)=M_1=\frac{\beta}{\alpha-1}$$

Multiplying (3.12b) by k^2 and sum the results over k

$$\sum_{k=1}^{\infty}k^2(\alpha+\beta+k)p_k=\sum_{k=1}^{\infty}k^2(\beta+k-1)p_{k-1}$$

$$(\alpha+\beta)\sum_{k=1}^{\infty}k^2p_k+\sum_{k=1}^{\infty}k^3p_k=\sum_{k=1}^{\infty}[(k-1)+1]^2[\beta+(k-1)]p_{k-1}$$

$$(\alpha+\beta)M_2+M_3=(\beta+2)\sum_{k=1}^{\infty}(k-1)^2p_{k-1}+(2\beta+1)\sum_{k=1}^{\infty}(k-1)p_{k-1}+\beta\sum_{k=1}^{\infty}p_{k-1}+\sum_{k=1}^{\infty}(k-1)^3p_{k-1}$$

$$\Longleftrightarrow (\alpha+\beta)M_2+M_3=(\beta+2)M_2+(2\beta+1)M_1+\beta+M_3$$

But $M_1=\frac{\beta}{\alpha-1}$
Thus

$$M_2=\frac{2\beta^2+\alpha\beta}{(\alpha-1)(\alpha-2)}$$

$$\begin{aligned}Var(X)&=M_2-[M_1]^2\\&=\frac{2\beta^2+\alpha\beta}{(\alpha-1)(\alpha-2)}-[\frac{\beta}{\alpha-1}]\\&=\frac{\alpha\beta(\alpha+\beta-1)}{(\alpha-1)^2(\alpha-2)}\end{aligned}$$

To obtain the pgf
Multiply (3.12b) by s^k and sum the results over k

$$\sum_{k=1}^{\infty}(\alpha+\beta+k)p_ks^k=\sum_{k=1}^{\infty}(\beta+k-1)p_{k-1}s^k$$

$$\Longleftrightarrow (\alpha+\beta)\sum_{k=1}^{\infty}p_ks^k+\sum_{k=1}^{\infty}kp_ks^k=\beta\sum_{k=1}^{\infty}p_{k-1}s^k+\sum_{k=1}^{\infty}(k-1)p_{k-1}s^k$$

i.e.;

$$(\alpha+\beta)\sum_{k=1}^{\infty}p_k s^k + s\sum_{k=1}^{\infty}kp_k s^{k-1} = \beta s\sum_{k=1}^{\infty}p_{k-1}s^{k-1} + s^2\sum_{k=1}^{\infty}(k-1)p_{k-1}s^{k-2}$$

$$\Longrightarrow (\alpha+\beta)[G(s)-p_0] + s\frac{dG}{ds} = \beta sG(s) + s^2\frac{dG}{ds}$$

But $p_0 = \frac{\Gamma(\alpha+1)}{(\alpha+\beta)}$
Therefore

$$\Longrightarrow s(1-s)\frac{dG}{ds} = (\beta s - \alpha - \beta)G(s) + \Gamma(\alpha+1)$$

which cannot be solved explicitly.

Type II:When X is the total number of trials required to achieve the first su

Given

$$p_k = \frac{\Gamma(\alpha+1)\Gamma(\alpha+\beta)\Gamma(\beta+k-1)}{\Gamma\beta\Gamma(\alpha+\beta+k)} \quad k=1,2,\ldots$$

$$p_1 = \frac{\Gamma(\alpha+1)}{(\alpha+\beta)}$$

Then

$$p_{k-1} = \frac{\Gamma(\alpha+1)\Gamma(\alpha+\beta)\Gamma(\beta+(k-1)-1)}{\Gamma\beta\Gamma(\alpha+\beta+(k-1))} \quad k=2,3,\ldots$$

$$\frac{p_k}{p_{k-1}} = \frac{\beta+k-2}{\alpha+\beta+k-1} \qquad (3.13a)$$

$$\Longrightarrow (\alpha+\beta+k-1)p_k = (\beta+k-2)p_{k-1} \qquad k=2,3,\ldots \qquad (3.13b)$$

Using Feller's method
Multiply (3.13b) by k and sum the results over k

$$\sum_{k=2}^{\infty}k(\alpha+\beta-1+k)p_k = \sum_{k=2}^{\infty}k(\beta+k-1)p_{k-1}$$

$$(\alpha+\beta-1)\sum_{k=2}^{\infty}kp_k + \sum_{k=2}^{\infty}k^2p_k = \sum_{k=2}^{\infty}[(k-1)+1][(\beta-1)+(k-1)]p_{k-1}$$

$$
\begin{aligned}
(\alpha+\beta-1)[M_1-p_1]+[M_2-p_1] &= [(\beta-1)]\sum_{k=2}^{\infty}(k-1)p_{k-1}+\sum_{k=2}^{\infty}(k-1)^2p_{k-1} \\
&\quad +(\beta-1)\sum_{k=2}^{\infty}p_{k-1}+\sum_{k=1}^{\infty}(k-1)p_{k-1}
\end{aligned}
$$

$$
\Longrightarrow (\alpha+\beta-1)[M_1-p_1]+[M_2-p_1]=(\beta-1)M_1+M_2+(\beta-1)+M_1
$$

But $p_1=\frac{\Gamma(\alpha+1)}{(\alpha+\beta)}$
Therefore this simplifies to

$$
\begin{aligned}
M_1(\alpha-1) &= (\alpha+\beta)p_1+(\beta-1) \\
&= \frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)}
\end{aligned}
$$

Therefore

$$
E(X)=M_1=\frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)}
$$

Multiply (3.13b) by k^2 and sum the results over k

$$
\sum_{k=2}^{\infty}k^2(\alpha+\beta-1+k)p_k=\sum_{k=2}^{\infty}[(k-1)+1]^2[(\beta-1)+(k-1)]p_{k-1}
$$

$$
(\alpha+\beta-1)\sum_{k=2}^{\infty}k^2p_k+\sum_{k=2}^{\infty}k^3p_k=\sum_{k=2}^{\infty}[(k-1)+1]^2[(\beta-1)+(k-1)]p_{k-1}
$$

$$
\begin{aligned}
(\alpha+\beta-1)[M_2-p_1]+[M_3-p_1] &= [(\beta-1)+2]\sum_{k=2}^{\infty}(k-1)^2p_{k-1} \\
&\quad +[2(\beta-1)+1]\sum_{k=2}^{\infty}(k-1)p_{k-1}+\sum_{k=2}^{\infty}(k-1)^3p_{k-1} \\
&\quad +(\beta-1)\sum_{k=2}^{\infty}p_{k-1}
\end{aligned}
$$

$$
\Longrightarrow (\alpha+\beta-1)[M_2-p_1]+[M_3-p_1]=(\beta+1)M_2+(2\beta-1)M_1+M_3+(\beta-1)
$$

$$
\Longleftrightarrow M_2(\alpha-2)=(\alpha+\beta)p_1+(2\beta-1)M_1+(\beta-1)
$$

Substituting $p_1 = \frac{\Gamma(\alpha+1)}{(\alpha+\beta)}$ and $M_1 = \frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)}$

$$M_2(\alpha-2) = (\alpha+\beta)\frac{\Gamma(\alpha+1)}{(\alpha+\beta)} + (2\beta-1)\frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)} + (\beta-1)$$
$$= \Gamma(\alpha+1) + (2\beta-1)\frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)} + (\beta-1)$$

$$M_2 = \frac{\Gamma(\alpha+1) + (2\beta-1)\frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)} + (\beta-1)}{(\alpha-2)}$$
$$= \frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)} + \frac{2\beta(\Gamma(\alpha+1)+\beta-1)}{(\alpha-1)(\alpha-2)}$$

Therefore

$$Var(X) = M_2 - [M_1]^2$$
$$= \frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)} + \frac{2\beta(\Gamma(\alpha+1)+\beta-1)}{(\alpha-1)(\alpha-2)} - [\frac{\Gamma(\alpha+1)+(\beta-1)}{(\alpha-1)}]^2$$

To obtain the pgf
Multiply (3.13b) by s^k and sum the results over k

$$\sum_{k=2}^{\infty}(\alpha+\beta-1+k)p_k s^k = \sum_{k=2}^{\infty}(\beta+k-1)p_{k-1}s^k$$

$$(\alpha+\beta-1)\sum_{k=2}^{\infty}p_k s^k + \sum_{k=2}^{\infty}kp_k s^k = (\beta-1)\sum_{k=2}^{\infty}p_{k-1}s^k + \sum_{k=2}^{\infty}(k-1)p_{k-1}s^k$$

$$\Longleftrightarrow (\alpha+\beta-1)\sum_{k=2}^{\infty}p_k s^k + s\sum_{k=2}^{\infty}kp_k s^{k-1} = (\beta-1)s\sum_{k=2}^{\infty}p_{k-1}s^k + s^2\sum_{k=2}^{\infty}(k-1)p_{k-1}s^{k-2}$$

$$\Longrightarrow (\alpha+\beta-1)[G(s)-p_0-p_1 s] + s[\frac{dG}{ds} - p_1] = (\beta-1)sG(s) + s^2\frac{dG}{ds}$$

But $p_0 = 0$ and $p_1 = \frac{\Gamma(\alpha+1)}{(\alpha+\beta)}$ therefore this simplifies to

$$s(1-s)\frac{dG}{ds} = \Gamma(\alpha+1) + [(\beta-1)s - (\alpha+\beta-1)]G(s)$$

which cannot be solved explicitly.

3.3.2 Beta - Binomial Distribution

Given

$$p_k = \binom{n}{k}\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+k)\Gamma(\beta+n-k)}{\Gamma\beta\Gamma\alpha\Gamma(\alpha+\beta+n)} \quad k=0,1,2,\ldots,n$$

then

$$p_{k-1} = \binom{n}{k-1}\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+(k-1))\Gamma(\beta+n-(k-1))}{\Gamma\beta\Gamma\alpha\Gamma(\alpha+\beta+n)} \quad k=1,2,\ldots,n$$

$$\frac{p_k}{p_{k-1}} = \frac{(n-k+1)(\alpha+k-1)}{k(n-k+\beta)} \qquad (3.14a)$$

$$k(n-k+\beta)p_k = (n-k+1)(\alpha+k-1)p_{k-1} \quad k=1,2,3,\ldots,n \qquad (3.14b)$$

Using Feller's method
Summing (3.14b) over k results to

$$\sum_{k=1}^{n} k(n-k+\beta)p_k = \sum_{k=1}^{n}[n-(k-1)][\alpha+(k-1)]p_{k-1}$$

$$(n+\beta)\sum_{k=1}^{n} kp_k = n\alpha\sum_{k=1}^{n} p_{k-1} + (n-\alpha)\sum_{k=1}^{n}(k-1)(p_{k-1} - \sum_{k=1}^{n}(k-1)^2 p_{k-1}$$

$$(n+\beta)M_1 = n\alpha + (n-\alpha)[M_1 - np_n] - [M_2 - n^2 p_n]$$

which simplifies to

$$(\alpha+\beta)M_1 = n\alpha$$

Therefore

$$E(X) = M_1 = \frac{n\alpha}{(\alpha+\beta)}$$

Multiply (3.14b) by k and sum the results over k

$$\sum_{k=1}^{n} k^2(n-k+\beta)p_k = \sum_{k=1}^{n}[(k-1)+1][n-(k-1)][\alpha+(k-1)]p_{k-1}$$

$$
\begin{aligned}
(n+\beta)\sum_{k=1}^{n} k^2 p_k + \sum_{k=1}^{n} k^3 p_k \;&=\; n\alpha[\sum_{k=1}^{n}(k-1)p_{k-1} + \sum_{k=1}^{n} p_{k-1}] \\
&\quad +(n-\alpha)[\sum_{k=1}^{n} k-1)^2 p_{k-1} + \sum_{k=1}^{n}(k-1)p_{k-1}] \\
&\quad -\sum_{k=1}^{n}(k-1)^3 p_{k-1} - \sum_{k=1}^{n}(k-1)^2 p_{k-1}
\end{aligned}
$$

$(n+\beta)M_2+M_3 = n\alpha[M_1-np_n]+(n-\alpha)[M_2-n^2p_n]-[M_3-n^3p_n]+n\alpha+(n-\alpha)[M_1-np_n]-[M_2$

which simplifies to

$$
(\alpha+\beta+1)M_2 = \frac{\alpha n(\alpha n+n)+\alpha\beta n}{(\alpha+\beta)}
$$

$$
\Longrightarrow M_2 = \frac{\alpha n(\alpha n+n)+\alpha\beta n}{(\alpha+\beta)(\alpha+\beta+1)}
$$

Therefore

$$
\begin{aligned}
Var(X) &= M_2 - [M_1]^2 \\
&= \frac{\alpha n(\alpha n+n)+\alpha\beta n}{(\alpha+\beta)(\alpha+\beta+1)} - [\frac{n\alpha}{(\alpha+\beta)}]^2 \\
&= \frac{\alpha\beta n(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}
\end{aligned}
$$

To obtain the pgf
Multiply (3.14b) by s^k and sum the results over k

$$
\sum_{k=1}^{n} k(n-k+\beta)p_k s^k = \sum_{k=1}^{n}[n-(k-1)][\alpha+(k-1)]]p_{k-1}s^k
$$

$$
(n+\beta)\sum_{k=1}^{n} kp_k s^k - \sum_{k=1}^{n} k^2 p_k s^k = n\alpha\sum_{k=1}^{n} p_{k-1}s^k + (n-\alpha)\sum_{k=1}^{n}(k-1)p_{k-1}s^k - \sum_{k=1}^{n}(k-1)^2 p_{k-1}s^k
$$

$\Longleftrightarrow$

$$
\begin{aligned}
(n+\beta)s\sum_{k=1}^{n} kp_k s^{k-1} - \sum_{k=1}^{n} k^2 p_k s^k \;&=\; n\alpha s\sum_{k=1}^{n} p_{k-1}s^{k-1} + (n-\alpha)s^2\sum_{k=1}^{n}(k-1)p_{k-1}s^{k-2} \\
&\quad -\sum_{k=1}^{n}(k-1)^2 p_{k-1}s^k
\end{aligned}
$$

$$\Longrightarrow (n+\beta)s\frac{dG}{ds}-s^2\frac{d^2G}{ds^2}-s\frac{dG}{ds} = n\alpha s[G(s)-p_ns^n]+(n-\alpha)s^2[\frac{dG}{ds}-np_ns^{n-1}]-s^3\frac{d^2G}{ds^2}-s^2\frac{dG}{ds}$$

which simplifies to

$$[(n+\beta-1)+(\alpha-n+1)s]\frac{dG}{ds}-s(1-s)\frac{d^2G}{ds^2} = n\alpha G(s)-n^2p_ns^n$$

which cannot be solved explicitly.

3.3.3 Beta - Negative Binomial Distribution

Type I: When X is the number of failures before the r^{th} success Given

$$p_k = \frac{\Gamma(k+r)\Gamma(\alpha+\beta)\Gamma(\alpha+r)\Gamma(\beta+k)}{\Gamma(k+1)\Gamma r\Gamma\alpha\Gamma\beta\Gamma(\alpha+\beta+k+r)}$$

then

$$p_{k-1} = \frac{\Gamma((k-1)+r)\Gamma(\alpha+\beta)\Gamma(\alpha+r)\Gamma(\beta+(k-1))}{\Gamma((k-1)+1)\Gamma r\Gamma\alpha\Gamma\beta\Gamma(\alpha+\beta+(k-1)+r)}$$

$$\frac{p_k}{p_{k-1}} = \frac{(k+r-1)(\beta+k-1)}{k(\alpha+\beta+k+r-1)} \qquad (3.15a)$$

$$k(\alpha+\beta+k+r-1)p_k = (k+r-1)(\beta+k-1)p_{k-1} \qquad (3.15b)$$

for $k = 1, 2, 3, \ldots$

Using Feller's method

Summing (3.15b) over k results to

$$\sum_{k=1}^{\infty} k(\alpha+\beta+k+r-1)p_k = \sum_{k=1}^{\infty}(k+r-1)(\beta+k-1)p_{k-1}$$

$$(\alpha+\beta+r-1)\sum_{k=1}^{\infty} kp_k + \sum_{k=1}^{\infty} k^2p_k = \sum_{k=1}^{\infty}[(k-1)+r][\beta+(k-1)]p_{k-1}$$

$$= (\beta+r)\sum_{k=1}^{\infty}(k-1)p_{k-1} + \sum_{k=1}^{\infty}(k-1)^2p_{k-1} + \beta r\sum_{k=1}^{\infty} p_{k-1}$$

$$\Longrightarrow (\alpha+\beta+r-1)M_1 + M_2 = (\beta+r)M_1 + M_2 + \beta r$$

which simplifies to

$$M_1 = \frac{\beta r}{\alpha - 1}$$

Therefore

$$E(X) = M_1 = \frac{\beta r}{\alpha - 1}$$

Multiply (3.15b) by k and sum the results over k

$$\sum_{k=1}^{\infty} k^2(\alpha + \beta + k + r - 1)p_k = \sum_{k=1}^{\infty} k(k + r - 1)(\beta + k - 1)p_{k-1}$$

$$(\alpha+\beta+r-1)\sum_{k=1}^{\infty} k^2 p_k + \sum_{k=1}^{\infty} k^3 p_k = \sum_{k=1}^{\infty}[(k-1)+1][(k-1)+r][\beta+(k-1)]p_{k-1}$$

i.e.;

$$\begin{aligned}(\alpha + \beta + r - 1)\sum_{k=1}^{\infty} k^2 p_k + \sum_{k=1}^{\infty} k^3 p_k &= (\beta + r)\sum_{k=1}^{\infty}(k-1)^2 p_{k-1} + \sum_{k=1}^{\infty}(k-1)^3 p_{k-1} \\ &\quad +(\beta r + \beta + r)\sum_{k=1}^{\infty}(k-1)p_{k-1} + \sum_{k=1}^{\infty}(k-1)^2 p_{k-1} \\ &\quad +\beta r\sum_{k=1}^{\infty} p_{k-1}\end{aligned}$$

$$\Longrightarrow (\alpha + \beta + r - 1)M_2 + M_3 = (\beta + r - 1)M_2 + M_3 + (\beta r + \beta + r)M_1 + \beta r$$

which simplifies to

$$M_2 = \frac{\beta r(\beta r + \beta + r + (\alpha - 1))}{\alpha(\alpha - 1)}$$

Therefore

$$\begin{aligned}Var(X) &= M_2 - [M_1]^2 \\ &= \frac{\beta r(\beta r + \beta + r + (\alpha - 1))}{\alpha(\alpha - 1)} - [\frac{\beta r}{\alpha - 1}]^2 \\ &= \frac{\beta^2 r + r^2\beta}{\alpha(\alpha - 1)} + \frac{r\beta}{\alpha} - \frac{\beta^2 r^2}{\alpha(\alpha - 1)^2}\end{aligned}$$

To obtain the pgf
Multiply (3.15b) by s^k and sum the results over k

$$\sum_{k=1}^{\infty} k(\alpha+\beta+k+r-1)p_k s^k = \sum_{k=1}^{\infty}(k+r-1)(\beta+k-1)p_{k-1}s^k$$

$$\begin{aligned}(\alpha+\beta+r-1)\sum_{k=1}^{\infty} kp_k s^k + \sum_{k=1}^{\infty} k^2 p_k s^k &= \sum_{k=1}^{\infty}[(k-1)+r][\beta+(k-1)]p_{k-1}s^k \\ &= (\beta+r)\sum_{k=1}^{\infty}(k-1)p_{k-1}s^k + \beta r\sum_{k=1}^{\infty} p_{k-1}s^k \\ &+ \sum_{k=1}^{\infty}(k-1)^2 p_{k-1}s^k\end{aligned}$$

$$\begin{aligned}(\alpha+\beta+r-1)s\sum_{k=1}^{\infty} kp_k s^{k-1} + \sum_{k=1}^{\infty} k^2 p_k s^k &= (\beta+r)s^2\sum_{k=1}^{\infty}(k-1)p_{k-1}s^{k-2} \\ &+ \beta rs\sum_{k=1}^{\infty} p_{k-1}s^{k-1} + \sum_{k=1}^{\infty}(k-1)^2 p_{k-1}s^k\end{aligned}$$

$$\Longleftrightarrow (\alpha+\beta+r-1)s\frac{dG}{ds}+s^2\frac{d^2G}{ds^2}+s\frac{dG}{ds} = (\beta+r)s^2\frac{dG}{ds}+\beta rsG(s)+s^3\frac{d^2G}{ds^2}+s^2\frac{dG}{ds}$$

$$\Longrightarrow [(\alpha+\beta+r)-(\beta+r-1)s]\frac{dG}{ds}+s(1-s)\frac{d^2G}{ds^2} = \beta rG(s)$$

which cannot be solved explicitly.

Type II: When X is the number of trials required to achieve the r^{th} success

Given

$$p_k = \frac{\Gamma(k)\Gamma(\alpha+\beta)\Gamma(\alpha+r)\Gamma(\beta+k-r)}{\Gamma(k+r+1)\Gamma r\Gamma\alpha\Gamma\beta\Gamma(\alpha+\beta+k)}$$

then

$$p_{k-1} = \frac{\Gamma(k-1)\Gamma(\alpha+\beta)\Gamma(\alpha+r)\Gamma(\beta+(k-1)-r)}{\Gamma((k-1)+r+1)\Gamma r\Gamma\alpha\Gamma\beta\Gamma(\alpha+\beta+(k-1))}$$

$$\frac{p_k}{p_{k-1}} = \frac{(k-1)(k+\beta-r-1)}{(k-r)(\alpha+\beta+k-1)} \qquad (3.16a)$$

$$(k-r)(\alpha+\beta+k-1)p_k = (k-1)(k+\beta-r-1)p_{k-1} \qquad (3.16b)$$

for $k = r+1, r+2, r+3, \ldots$

Using Feller's method

Multiply (3.16b) by k and sum the results over k

$$\sum_{k=r+1}^{\infty} k(k-r)(\alpha+\beta+k-1)p_k = \sum_{k=r+1}^{\infty} k(k-1)(k+\beta-r-1)p_{k-1}$$

$$\begin{aligned}(\alpha+\beta-r-1)[\sum_{k=r+1}^{\infty} k^2p_k - r\sum_{k=r+1}^{\infty} kp_k] + \sum_{k=r+1}^{\infty} k^3p_k &= (\beta-r+1)\sum_{k=r+1}^{\infty}(k-1)^2p_{k-1} \\ &+ \sum_{k=r+1}^{\infty}(k-1)^3p_{k-1} + \sum_{k=r+1}^{\infty}(k- \\ &-(\beta-r)\sum_{k=r+1}^{\infty}(k-1)p_{k-1}\end{aligned}$$

$$\Longrightarrow (\alpha+\beta-r-1)M_1 + M_3 - r(\alpha+\beta-1)M_1 - rM_2 = (\beta-r)M_1 - rM_2 + M_3$$

which simplifies to

$$M_1 = \frac{r(\alpha+\beta-1)}{(\alpha-1)}$$

Therefore

$$E(X) = M_1 = \frac{r(\alpha+\beta-1)}{(\alpha-1)}$$

Multiply (3.16b) by k^2 and sum the results over k

$$\sum_{k=r+1}^{\infty} k^2(k-r)(\alpha+\beta+k-1)p_k = \sum_{k=r+1}^{\infty} k^2(k-1)(k+\beta-r-1)p_{k-1}$$

$$\begin{aligned}(\alpha+\beta-r-1)[\sum_{k=r+1}^{\infty} k^2p_k - r\sum_{k=r+1}^{\infty} kp_k] + \sum_{k=r+1}^{\infty} k^3p_k &= (\beta-r+1)\sum_{k=r+1}^{\infty}(k-1)^2p_{k-1} \\ + \sum_{k=r+1}^{\infty}(k-1)^3p_{k-1} + (\beta-r)\sum_{k=r+1}^{\infty}(k-1)p_{k-1}\end{aligned}$$

$$\Longrightarrow (\alpha+\beta-r-1)M_2 + M_3 + (r\alpha+r\beta-r) = (\beta-r)M_2 + M_3 + (\beta-r)M_1 + M_2$$

which simplifies to

$$M_2 = \frac{r(r\beta + r\alpha + \beta)(\alpha + \beta - 1)}{(\alpha - 1)(\alpha - 2)}$$

Therefore

$$\begin{aligned} Var(X) &= M_2 - [M_1]^2 \\ &= \frac{r(r\beta + r\alpha + \beta)(\alpha + \beta - 1)}{(\alpha - 1)(\alpha - 2)} - [\frac{r(\alpha + \beta - 1)}{(\alpha - 1)}]^2 \\ &= \frac{r\alpha\beta}{(\alpha - 1)^2} + \frac{2r^2 + r\beta^2}{(\alpha - 1)(\alpha - 2)} + r\beta(r\alpha + r\beta + r + 1)(\alpha - 1)^2(\alpha - 2) \end{aligned}$$

To obtain the pgf
Multiply (3.16b) by s^k and sum the results over k

$$\sum_{k=r+1}^{\infty} (k - r)(\alpha + \beta + k - 1)p_k s^k = \sum_{k=r+1}^{\infty} (k - 1)(k + \beta - r - 1)p_{k-1} s^k$$

$$(\alpha + \beta - r - 1) \sum_{k=r+1}^{\infty} kp_k s^k + \sum_{k=r+1}^{\infty} k^2 p_k s^k - r(\alpha + \beta - 1) \sum_{k=r+1}^{\infty} p_k s^k =$$
$$(\beta - r) \sum_{k=r+1}^{\infty} (k - 1)p_{k-1} s^k + \sum_{k=r+1}^{\infty} (k - 1)^2 p_{k-1} s^k$$

which is equivalent to

$$(\alpha + \beta - r - 1)s \sum_{k=r+1}^{\infty} kp_k s^{k-1} + \sum_{k=r+1}^{\infty} k^2 p_k s^k - r(\alpha + \beta - 1) \sum_{k=r+1}^{\infty} p_k s^k =$$
$$(\beta - r)s^2 \sum_{k=r+1}^{\infty} (k - 1)p_{k-1} s^{k-2} + \sum_{k=r+1}^{\infty} (k - 1)^2 p_{k-1} s^k$$

This implies that

$$(\alpha + \beta - r - 1)s[\frac{dG}{ds} - rp_r s^{r-1}] + s^2\frac{d^2G}{ds^2} + s\frac{dG}{ds} - r(\alpha + \beta - 1)[G(s) - p_r s^r] =$$
$$(\beta - r)s^2\frac{dG}{ds} + s^3\frac{d^2G}{ds^2} + s^2\frac{dG}{ds}$$

$$\Longleftrightarrow [(\alpha+\beta-r)s-(\beta-r+1)s^2]\frac{dG}{ds} + s^2(1-s)\frac{d^2G}{ds^2} = (\alpha+\beta-1)rG(s) - r^2 p_r s^r$$

which cannot be solved explicitly.

Chapter 4

PATTERNS OF RECURSIVE RELATIONS AND THEIR CHARACTERIZATIONS

4.1 Introduction

Our interest in this chapter is to identify patterns of recursive relations that were derived in chapter 3.

For each pattern,examples of probability distributions satisfying it are given.

An attempt has also been made to study some properties of particular patterns.

Research questions to be addressed are:

1) Given a recursive pattern,can one identify probability distributions satisfying it?

2) Can we find the pgf,mean and variance of the recursive relation with that particular pattern?

3) What are the estimators of the parameters of the recurrence equation with that particular pattern?

4.2 Patterns of Recursive Relations

4.2.1 Panjer's Pattern (1981)

$$P_n = (a + \frac{b}{n})P_{n-1} \qquad n = 1, 2, \ldots$$

Examples of distributions satisfying this recursive pattern are:
i)Poisson Distribution

$$P_n = \frac{\lambda}{n}P_{n-1} \qquad n = 1, 2, \ldots$$

$\Longrightarrow a = 0$ and $b = \lambda$
ii)Binomial Distribution

$$P_n = \frac{(m-n+1)p}{qn}p_{n-1} \qquad n = 1, 2, \ldots, m$$

$$= [-\frac{p}{q} + \frac{(m+1)p}{qn}]P_{n-1}$$

$\Longrightarrow a = -\frac{p}{q}$ and $b = \frac{(m+1)p}{q}$
iii)Negative Binomial Distribution Type I
When n is the number of failures before the r^{th} success

$$P_n = q\frac{n+r-1}{n}P_{n-1} \qquad n = 1, 2, 3, \ldots$$

where $q = 1 - p$

$$= [q + \frac{q(r-1)}{n}]P_{n-1}$$

$\Longrightarrow a = q$ and $b = q(r-1)$

4.2.2 Willmot's pattern (1988)

$$P_n = (a + \frac{b}{n})P_{n-1} \qquad n = 2, 3, \ldots$$

Examples of distributions satisfying this recursive pattern are:
i)Logarithmic Distribution

$$P_n = \frac{p(n-1)}{n} P_{n-1} \qquad n = 2, 3, \ldots$$

$$= [p + \frac{-p}{n}] P_{n-1}$$

$\Longrightarrow a = p$ **and** $b = -p$

ii)Geometric Distribution Type II

When n is the total number of trials required to achieve the first success

$$P_n = qP_{n-1} \qquad n = 2, 3, \ldots$$

$\Longrightarrow a = p$ **and** $b = 0$

iii)Zero-Truncated Poisson Distribution

$$P_n = \frac{\lambda}{n} P_{n-1} \qquad n = 2, 3, 4, \ldots$$

$\Longrightarrow a = 0$ **and** $b = \lambda$

iv)Zero-Truncated Negative Binomial Distribution

$$P_n = q\frac{n+r-1}{n} P_{n-1} \qquad n = 2, 3, 4, \ldots$$

where $q = 1 - p$

$$= [q + \frac{q(r-1)}{n}] P_{n-1}$$

$\Longrightarrow a = q$ **and** $b = q(r-1)$

4.2.3 Schroter(1990)

$$P_n = (a + \frac{b}{n}) P_{n-1} + \frac{c}{n} P_{n-2} \qquad n = 1, 2, \ldots$$

No examples of probability distributions satisfying this pattern were found.

4.2.4 Sundt(1992)

$$P_n = \sum_{j=1}^{k}(a_j + \frac{b_j}{n})P_{n-1} \qquad n = \omega+1, \omega+2, \ldots$$

Special cases

i)When $k = 1$ and $\omega = 0$ we have

4.2.4.1 Panjer's Pattern

$$P_n = (a + \frac{b}{n})P_{n-1} \qquad n = 1, 2, \ldots$$

which is the Panjer(1981) pattern.

ii)When $k = 1$ and $\omega = 1$ we have

4.2.4.2 The zero-truncated pattern

$$P_n = (a + \frac{b}{n})P_{n-1} \qquad n = 2, 3, \ldots$$

iii)When $k = 2$ and $\omega = 0$ we have

$$P_n = \sum_{j=1}^{2}(a_j + \frac{b_j}{n})P_{n-1} \quad ; n = 1, 2, \ldots$$

i.e.

$$P_n = [(a_1 + \frac{b_1}{n}) + (a_2 + \frac{b_2}{n})]P_{n-1} \quad ; n = 1, 2, \ldots$$

No examples of probability distributions satisfying this case were found.

iv)When $k = 4$ and $\omega = 0$ we have

i)Hyper-Geometric Distribution

$$P_y = \frac{(d-y+1)(x-y+1)}{n(N-d-x+y)}P_{y-1} \qquad y = 1, 2, \ldots, d$$

$$= [\frac{(y-d-x-2}{(N-d-x+y)} + \frac{(dx+d+x+1)}{y(N-d-x+y)}]p_{y-1}$$

$k = 4$ and

$a_1 = \frac{y}{(N-d-x+y)}$ $a_2 = \frac{-d}{(N-d-x+y)}$ $a_3 = \frac{-x}{(N-d-x+y)}$ $a_4 = \frac{-2}{(N-d-x+y)}$

$b_1 = \frac{dx}{(N-d-x+y)}$ $b_2 = \frac{d}{(N-d-x+y)}$ $b_3 = \frac{x}{(N-d-x+y)}$ $b_4 = \frac{1}{(N-d-x+y)}$

ii)Beta-Binomial Distribution

$$P_n = \frac{(m-n+1)(\alpha+n-1)}{n(m-n+\beta)}P_{n-1} \qquad n = 1, 2, \ldots, m$$

$$= [\frac{(m-\alpha-n+2)}{(m-n+\beta)} + \frac{(m\alpha+\alpha-m-1)}{n(m-n+\beta)}]P_{n-1}$$

$k = 4$ and

$a_1 = \frac{m}{(m-n+\beta)}$ $a_2 = \frac{-\alpha}{(m-n+\beta)}$ $a_3 = \frac{-n}{(m-n+\beta)}$ $a_4 = \frac{2}{(m-n+\beta)}$

$b_1 = \frac{-m}{(m-n+\beta)}$ $b_2 = \frac{\alpha}{(m-n+\beta)}$ $b_3 = \frac{m\alpha}{(m-n+\beta)}$ $b_4 = \frac{-1}{(m-n+\beta)}$

iii)Beta-Negative Binomial Distribution

Case1:When n is the number of failures before the r^{th} success

$$P_n = q\frac{(n+r-1)(\beta+n-1)}{n(\alpha+\beta+n+r-1)}P_{n-1} \qquad n = 1, 2, 3, \ldots$$

$$= [\frac{(\beta+n+r-2)}{(\alpha+\beta+n+r-1)} + \frac{r\beta-r-\beta+1}{n(\alpha+\beta+n+r-1)}]P_{n-1}$$

$k = 4$ and

$a_1 = \frac{\beta}{(\alpha+\beta+n+r-1)}$ $a_2 = \frac{r}{(\alpha+\beta+n+r-1)}$ $a_3 = \frac{n}{(\alpha+\beta+n+r-1)}$ $a_4 = \frac{-2}{(\alpha+\beta+n+r-1)}$

$b_1 = \frac{-\beta}{(\alpha+\beta+n+r-1)}$ $b_2 = \frac{-r}{(\alpha+\beta+n+r-1)}$ $b_3 = \frac{r\beta}{(\alpha+\beta+n+r-1)}$ $b_4 = \frac{1}{(\alpha+\beta+n+r-1)}$

4.2.5 Panjer and Willmot(1982)

$$P_n = \frac{\sum_{t=0}^{k} a_t n^t}{\sum_{t=0}^{k} b_t n^t}P_{n-1} \qquad n = 2, 3, \ldots, P_0 = 0$$

No examples of probability distributions satisfying this pattern were found.

4.2.6 Panjer and Sundt(1982)

$$P_n = \frac{\beta_0 + \beta_1(n-1) + \beta_2(n-1)^{(2)} + \ldots}{\alpha_0 + \alpha_1 n + \alpha_2 n^{(2)} + \ldots}P_{n-1} \qquad n = 1, 2, \ldots$$

where $n^{(k)} = n(n-1)(n-2)\ldots(n-k+1)$

Examples of distributions satisfying this recursive pattern are:

i)Poisson Distribution

$$P_n = \frac{\lambda}{n} P_{n-1} \qquad n = 1, 2, \ldots$$

$\Longrightarrow k = 1 \ \beta_0 = \lambda \ \beta_1 = 0 \ \alpha_0 = 0 \ \alpha_1 = 1$

ii)Binomial Distribution

$$P_n = \frac{(m-n+1)p}{qn} P_{n-1} \qquad n = 1, 2, \ldots, m$$

$$= [-\frac{p}{q} + \frac{(m+1)p}{qn}]p_{n-1}$$

$\Longrightarrow k = 1 \ \beta_0 = mp \ \beta_1 = -p \ \alpha_0 = 0 \ \alpha_1 = q$

iii)Negative Binomial Distribution

Case1:When n is the number of failures before the r^{th} success

$$P_n = q\frac{n+r-1}{n} P_{n-1} \qquad n = 1, 2, 3, \ldots$$

where $q = 1 - p$

$$= [q + \frac{q(r-1)}{n}]P_{n-1}$$

$\Longrightarrow k = 1 \ \beta_0 = qr \ \beta_1 = q \ \alpha_0 = 0 \ \alpha_1 = 1$

iv)Beta-Geometric Distribution

Case1:When n is the number of failures before the first success

$$P_n = \frac{\beta + (n-1)}{(\alpha+\beta) + n} P_{n-1} \qquad n = 1, 2, \ldots$$

$\Longrightarrow k = 1 \ \beta_0 = \beta \ \beta_1 = 1 \ \alpha_0 = \alpha + \beta \ \alpha_1 = 1$

4.3 Characterization of Recursive Patterns

4.3.1 Original Panjer's Pattern

Let

$$P_n = (a + \frac{b}{n})P_{n-1} \ ; n = 1, 2, \ldots$$

and $P_0 > 0$
Therefore

$$\begin{aligned} nP_n &= (an+b)P_{n-1} \\ &= [a(n-1+1)+b]P_{n-1} \\ &= a(n-1)P_{n-1} + (a+b)P_{n-1} \ ; n = 1, 2, \ldots \end{aligned}$$

Thus

$$\sum_{n=1}^{\infty} nP_n s^n = a\sum_{n=1}^{\infty}(n-1)P_{n-1}s^n + (a+b)\sum_{n=1}^{\infty} P_{n-1}s^n$$

$$\Longrightarrow s\frac{dG}{ds} = as^2\sum_{n=1}^{\infty}(n-1)P_{n-1}s^{n-2} + (a+b)s\sum_{n=1}^{\infty} P_{n-1}s^{n-1}$$

Thus

$$\frac{dG}{ds} = as\frac{dG}{ds} + (a+b)G(s)$$

Therefore

$$(1-as)\frac{dG}{ds} = (a+b)G(s) \tag{4.1}$$

Solving (4.1)
Case (i):$a = 0$,$b = 0$

$$\frac{dG}{ds} = 0 \Longrightarrow G(s) = c \Longrightarrow 1 = G(1) = c$$

i.e.

$$G(s) = 1$$

Therefore

$$P_n = \begin{cases} 1 & for \ \ n=0 \\ 0 & for \ \ n>0 \end{cases}$$

Case (ii):$a = 0$,$b \neq 0$

$$\frac{dG}{ds} = bG(s) \Longrightarrow \frac{1}{G(s)}\frac{dG}{ds} = b$$

$$\frac{d}{ds}logG(s) = b \Longrightarrow logG(s) = bs + c$$

Thus

$$G(s) = c_1e^{bs} \Longrightarrow 1 = G(1) = c_1e^{b}$$

implyng that

$$c_1 = e^{-b}$$

Therefore

$$G(s) = e^{-b(1-s)}$$

which is the pgf of a Poisson distribution with parameter b i.e.

$$P_n = \frac{e^{-b}b^n}{n!} \ for \ \ n = 0, 1, 2, \ldots \qquad (b > 0)$$

<u>Case (iii):$a \neq 0, b = 0$</u>

$$(1 - as)\frac{dG}{ds} = aG(s)$$

$$\frac{1}{G(s)}\frac{dG}{ds} = \frac{a}{1 - as} \Longrightarrow \frac{d}{ds}logG(s) = \frac{a}{1 - as}$$

$$logG(s) = -\int -\frac{a}{1 - as}ds = -logc(1 - as)$$

Therefore

$$G(s) = \frac{c_1}{1 - as} \Longrightarrow 1 = G(1) = \frac{c_1}{1 - as}$$

$$\Longrightarrow c_1 = 1 - a$$

Thus

$$G(s) = \frac{1 - a}{1 - as}$$

which is the pgf of a Geometric distribution with probability 1-a i.e.

$$P_n = a^n(1 - a) \ \ n = 0, 1, 2, \ldots$$

<u>Case (iv):$a \neq 0, b \neq 0$</u>

$$(1 - as)\frac{dG}{ds} = (a + b)G(s)$$

$a \neq 0, b \neq 0, a + b = 0$ we go back to case(i)
$a \neq 0, b \neq 0, a + b > 0$
Therefore

$$\frac{1}{G(s)}\frac{dG}{ds} = \frac{a + b}{1 - as}$$

$$\implies \frac{d}{ds}logG(s) = (a+b)\int \frac{ds}{1-as} = \frac{(a+b)}{-a}\int \frac{-a}{1-as}ds$$

Thus

$$\begin{aligned} logG(s) &= \frac{(a+b)}{-a}logc(1-as) \\ &= -mlogc(1-as) \end{aligned}$$

where $m = \frac{(a+b)}{a}$ implying that

$$G(s) = c_1(1-as)^{-m}$$

$$1 = G(1) = c_1(1-a)^{-m}$$

$$\implies c_1 = (1-a)^m$$

Hence

$$G(s) = (\frac{1-a}{1-as})^m$$

If m is a positive integer,then

$$\begin{aligned} G(s) &= (1-a)^m(1-as)^{-m} \\ &= (1-a)^m \sum_{n=0}^{\infty} \binom{-m}{n}(-as)^n \\ &= (1-a)^m \sum_{n=0}^{\infty} (-1)^n \binom{-m}{n} a^n s^n \\ &= (1-a)^m \sum_{n=0}^{\infty} \binom{m+n-1}{n} a^n s^n \end{aligned}$$

Therefore

$$P_n = \binom{m+n-1}{n} a^n (1-a)^m \;\; for \;\; 0 < a < 1, m > 0$$

If m is a negative integer,let $m = -\alpha$ where α is a positive integer.

$$\begin{aligned} G(s) &= (1-a)^{-\alpha}(1-as)^{\alpha} \\ &= (1-a)^{-\alpha} \sum_{n=0}^{\alpha} \binom{\alpha}{n}(-as)^n \end{aligned}$$

that is

$$G(s) = \sum_{n=0}^{\alpha} \binom{\alpha}{n} \frac{(-as)^n}{(1-a)^\alpha}$$

$$= \sum_{n=0}^{\alpha} \binom{\alpha}{n} \frac{(-as)^n}{(1-a)^n(1-a)^{\alpha-n}}$$

$$= \sum_{n=0}^{\alpha} \binom{\alpha}{n} (\frac{-a}{1-a})^n (\frac{1}{1-a})^{\alpha-n} s^n$$

Let

$$p = -\frac{a}{1-a}$$

then

$$1 - p = 1 + \frac{a}{1-a} = \frac{1}{1-a}$$

Therefore

$$G(s) = \sum_{n=0}^{\alpha} \binom{\alpha}{n} p^n (1-p)^{\alpha-n} s^n \;\; ; a < 0$$

implying that

$$P_n = \binom{\alpha}{n} p^n (1-p)^{\alpha-n} \;\; for \;\; n = 0, 1, 2, \ldots, \alpha$$

Thus we have the following theorem:

Theorem:

The recursive relation

$$P_n = (a + \frac{b}{n}) P_{n-1} \;\; ; n = 1, 2, \ldots$$

and $P_0 > 0$
satisfies

$$(1 - as)\frac{dG}{ds} = (a+b)G(s)$$

which gives

(i)

$$P_n = \begin{cases} 1 & for \;\; n=0 \\ 0 & for \;\; n>0 \end{cases} \qquad a = 0, b = 0$$

(ii)

$$P_n = \frac{e^{-b}b^n}{n!} \ for \ n = 0, 1, 2, \ldots \qquad b > 0 \ and \ a = o$$

(iii)

$$P_n = a^n(1-a) \ n = 0, 1, 2, \ldots \qquad 0 < a < 1 \ and \ b = 0$$

(iv)

$$P_n = \binom{m+n-1}{n} a^n(1-a)^m \ for \ 0 < a < 1, m = \frac{(a+b)}{a} > 0$$

(v)

$$P_n = \binom{\alpha}{n} p^n(1-p)^{\alpha-n} \ ; ln = 0, 1, 2, \ldots, \alpha \qquad p = -\frac{a}{1-a}, a < 0$$

So given

$$P_n = (a + \frac{b}{n})P_{n-1} \ ; n = 1, 2, \ldots$$

we have found that when
$a < 0, P_n$ **is a Binomial distribution**
$a = 0, P_n$ **is a Poisson distribution**
$a > 0, P_n$ **is a Negative Binomial distribution**

4.3.1.1 Applying the PGF technique on the original Panjer pattern

$$P_n \ = \ (a + \frac{b}{n})P_{n-1} \ for \ n = 1, 2, \ldots \tag{4.2}$$

a) Special cases of a and b.
Let us consider equation(1) for
Case 1:When $a = 0$ **and** $b \neq 0$
$P_n = \frac{b}{n} P_{n-1}$ **for** $n = 1, 2, \ldots$

$$\Longrightarrow nP_n = bP_{n-1}$$

Multiplying by s^n **and summing up over n results to**

$$\sum_{n=1}^{\infty} nP_n s^n = b\sum_{n=1}^{\infty} P_{n-1}s^n$$

$$\iff s\sum_{n=1}^{\infty} nP_n s^{n-1} = bs\sum_{n=1}^{\infty} P_{n-1}s^{n-1}\ldots(*)$$

Recal:

$$G(s) = \sum_{n=0}^{\infty} P_n s^n = \sum_{n=1}^{\infty} P_{n-1}s^{n-1}$$

$$G'(s) = \frac{dG}{ds} = \sum_{n=1}^{\infty} nP_n s^{n-1} = \sum_{n=1}^{\infty}(n-1)P_{n-1}s^{n-2}$$

Therefore equation $(*)$ becomes

$$s\frac{dG}{ds} = bsG(s)$$

$$\Longrightarrow \frac{dG}{ds} = bG(s)$$

Thus

$$\int \frac{dG}{G} = b\int ds$$

$$i.e ln[G(s)] = bs + C$$

$$\Longrightarrow G(s) = e^{bs+C} = e^{bs}e^{C} = ke^{bs}$$

When $s = 1$,then

$$G(1) = ke^{b}$$

$$\iff 1 = ke^{b}$$

$$\Longrightarrow k = e^{-b}$$

Hence

$$G(s) = e^{-b}e^{bs}$$

$$G(s) = e^{-b(1-s)} \qquad (4.3)$$

which is the pgf of a Poisson distribution with parameter b i.e.

$$P_n = \frac{e^{-b}b^n}{n!} \quad for \ n = 0, 1, 2, \ldots \qquad (4.4)$$

Case 2:When $a \neq 0$ and $b = 0$

$$P_n = aP_{n-1} \qquad \text{for } n = 1, 2, \ldots$$

$$\Longrightarrow nP_n = anP_{n-1}$$

Multiplying by s^n and summing up over n results to

$$\sum_{n=1}^{\infty} nP_n s^n = a \sum_{n=1}^{\infty} nP_{n-1} s^n$$

$$\Longleftrightarrow s \sum_{n=1}^{\infty} nP_n s^{n-1} = a \sum_{n=1}^{\infty} [(n-1)+1] P_{n-1} s^n$$

$$= a[\sum_{n=1}^{\infty} (n-1) P_{n-1} s^n + \sum_{n=1}^{\infty} P_{n-1} s^n]$$

$$= as^2 \sum_{n=1}^{\infty} (n-1) P_{n-1} s^{n-2} + as \sum_{n=1}^{\infty} P_{n-1} s^{n-1}$$

$$\Longrightarrow s\frac{dG}{ds} = as^2 \frac{dG}{ds} + asG(s)$$

$$\Longleftrightarrow s(1-as)\frac{dG}{ds} = asG(s)$$

$$i.e. (1-as)\frac{dG}{ds} = aG(s)$$

Therefore

$$\int \frac{dG}{G} = \int \frac{ads}{(1-as)}$$

$$lnG(s) = -ln(1-as) + lnk$$

$$\Longrightarrow G(s) = \frac{k}{(1-as)}$$

$$G(1) = \frac{k}{(1-a)}$$

$$\Longrightarrow k = 1 - a$$

since $G(1) = 1$
Therefore

$$G(s) = \frac{1-a}{(1-as)} \tag{4.5}$$

which is a probability generating function of Geometric Type I disribution.

Case 3:When $a \neq 0$ and $b \neq 0$

$P_n = (a + \frac{b}{n})P_{n-1} \qquad n = 1, 2, \ldots,$

$$\Longleftrightarrow P_n = (\frac{an + b}{n})P_{n-1}$$

$$nP_n = (an + b)P_{n-1}$$

Multiplying by s^n and summing up over n results to

$$\sum_{n=1}^{\infty} nP_n s^n = \sum_{n=1}^{\infty}(an + b)P_{n-1}s^n$$

$$s\sum_{n=1}^{\infty} nP_n s^{n-1} = a\sum_{n=1}^{\infty} nP_{n-1}s^n + b\sum_{n=1}^{\infty} P_{n-1}s^n$$

$$= a\sum_{n=1}^{\infty}[(n-1) + 1]P_{n-1}s^n + b\sum_{n=1}^{\infty} P_{n-1}s^n$$

$$= a[\sum_{n=1}^{\infty}(n-1)P_{n-1}s^n + \sum_{n=1}^{\infty} P_{n-1}s^n] + b\sum_{n=1}^{\infty} P_{n-1}s^n$$

$$= as^2\sum_{n=1}^{\infty}(n-1)P_{n-1}s^{n-2} + as\sum_{n=1}^{\infty} P_{n-1}s^{n-1} + bs\sum_{n=1}^{\infty} P_{n-1}s^{n-1}$$

$$\Longrightarrow s\frac{dG}{ds} = as^2\frac{dG}{ds} + (a + b)sG(s)$$

$$\Longleftrightarrow s(1 - as)\frac{dG}{ds} = (a + b)sG(s)$$

$$i.e.(1 - as)\frac{dG}{ds} = (a + b)G(s)$$

Therefore

$$\int \frac{dG}{G} = \int (a + b)\frac{ds}{(1 - as)}$$

$$lnG(s) = -\frac{(a + b)}{a}ln(1 - as) + C = -\frac{(a + b)}{a}ln(1 - as) + lnk$$

$$\Longrightarrow G(s) = k(1 - as)^{-\frac{(a+b)}{a}}$$

When $s = 1$,**then**

$$G(1) = k(1-a)^{-\frac{(a+b)}{a}}$$

$$i.e.1 = k(1-a)^{-\frac{(a+b)}{a}}$$

$$\Longrightarrow k = (1-a)^{\frac{(a+b)}{a}}$$

Thus substituting this value of k,

$$G(s) = (1-a)^{\frac{(a+b)}{a}}(1-as)^{-\frac{(a+b)}{a}}$$

Let $a + b = \alpha$ **and** $a = \beta$

$$G(s) = (1-\beta)^{\frac{\alpha}{\beta}}(1-\beta s)^{\frac{-\alpha}{\beta}} \quad (4.6)$$

There are now two subcases to be considered.

Sub-Case 1: When $0 < \beta < 1$

Let $\frac{\alpha}{\beta} = r$ **and** $1 - \beta = p$.**Then equation(4.5) becomes**

$$G(s) = p^r(1-qs)^{-r} = (\frac{p}{1-qs})^r \quad (4.7)$$

where $q = 1 - p = \beta$.

This is the pgf of a Negative Binomial distribution.

Note:

$$G(s) = p^r(1-qs)^{-r}$$

$$= p^r\sum_{k=0}^{\infty}\binom{-r}{n}(-qs)^n$$

$$= p^r\sum_{k=0}^{\infty}(-1)^n\binom{-r}{n}(qs)^n$$

$$= p^r\sum_{k=0}^{\infty}\binom{r+n-1}{n}q^n s^n$$

Since

$$(-1)^n\binom{-r}{n} = \binom{r+n-1}{n}$$

Therefore

$$P_n = \binom{r+n-1}{n}q^n p^r \;\; n = 0, 1, 2, \ldots \quad (4.8)$$

Sub-Case 2: When $\beta < 0$
Let $\frac{\alpha}{\beta} = \nu$ where $\nu > 0$.Then equation(4.5) becomes

$$
\begin{aligned}
G(s) &= (1-\beta)^{-\nu}(1-\beta s)^{\nu} \\
&= (\frac{1-\beta}{1-\beta s})^{\nu} \\
&= (\frac{1}{1-\beta} - \frac{\beta}{1-\beta}s)^{\nu} \\
&= (p+qs)^{\nu}
\end{aligned}
$$

where $p = \frac{1}{1-\beta}$ and $q = 1 - p = \frac{-\beta}{1-\beta}$
Therefore G(s) is the pgf of a binomial distribution with parameters p and ν i.e.

$$P_n = \binom{\nu}{n} p^n q^{\nu-n} \ for \ n = 0, 1, 2, \ldots, \nu \tag{4.9}$$

4.3.1.2 Applying the Pearson Difference (Iteration) technique on the original Panjer pattern

Let in general,

$$\frac{f(x+1)}{f(x)} = \frac{P(x)}{Q(x)}$$

where
f(.) is a discrete probability distribution;
P(x) and Q(x) are polynomials.
In particular,suppose
$P(x) = \alpha + \beta x$ **and** $Q(x) = x + 1$
Then

$$\frac{f(x+1)}{f(x)} = \frac{\alpha + \beta x}{x+1}$$

$\Longrightarrow$

$$f(x+1) = \frac{\alpha + \beta x}{x+1} f(x) \ for \ x = 0, 1, 2, 3, \ldots \tag{4.10}$$

where

$$f(x) = Prob(X = x) \geq 0 \tag{4.11}$$

and

$$\sum_{x=0}^{\infty} f(x) = 1 \tag{4.12}$$

Equation(4.9) implies

$$f(x) = \frac{\alpha + \beta(x-1)}{x} f(x-1) \; ; x = 1, 2, 3, \ldots \tag{4.13}$$

<u>Note</u>

Equation(4.12) is similar to the recursive pattern according to Panjer (1981) i.e.

$$\begin{aligned} P_n &= (a + \frac{b}{n}) P_{n-1} \; ; n = 1, 2, 3, \ldots \\ &= \frac{an + b}{n} P_{n-1} \\ &= \frac{b + a(n - 1 + 1)}{n} P_{n-1} \\ &= \frac{(a+b) + a(n-1)}{n} P_{n-1} \\ &= \frac{\alpha + \beta(n-1)}{n} P_{n-1} \end{aligned}$$

where $\alpha = a + b$ and $\beta = a$

Let us now consider equation(4.12) for $\beta = 0$ and $\beta \neq 0$

Case 1: When $\beta = 0$

$$f(x) = \frac{\alpha}{x} f(x-1) \; ; x = 1, 2, 3, \ldots \tag{4.14}$$

Therefore

$$\begin{aligned} x = 1 &\Longrightarrow f(1) = \frac{\alpha}{1} f(0) \\ x = 2 &\Longrightarrow f(2) = \frac{\alpha}{2} f(1) = \frac{\alpha^2}{2!} f(0) \\ x = 3 &\Longrightarrow f(3) = \frac{\alpha}{3} f(2) = \frac{\alpha^3}{3!} f(0) \end{aligned}$$

In general,

$$f(x) = \frac{\alpha^x}{x!} f(0); x = 1, 2, 3, \ldots$$

$\Longrightarrow$

$$f(x) = \frac{\alpha^x}{x!} f(0) \; ; x = 0, 1, 2, \ldots \qquad (4.15)$$

$$\Longrightarrow \sum_{x=0}^{\infty} f(x) = f(0) \sum_{x=0}^{\infty} \frac{\alpha^x}{x!} = f(0) e^{\alpha}$$

Therefore

$$1 = f(0) e^{\alpha}$$

$\Longrightarrow$

$$f(0) = e^{-\alpha} \qquad (4.16)$$

Thus

$$f(x) = \frac{e^{-\alpha} \alpha^x}{x!} \; for \; x = 0, 1, 2, \ldots \qquad (4.17)$$

which is a Poisson mass function with parameter α.

Case 2: When $\beta \neq 0$

Then equation(4.12) can be re-written as

$$\begin{aligned} f(x) &= \frac{\alpha + \beta(x-1)}{x} f(x-1) \\ &= \frac{\beta}{x} [\frac{\alpha + \beta(x-1)}{\beta}] f(x-1) \\ &= \frac{\beta}{x} [\frac{\alpha}{\beta} + x - 1] f(x-1) \end{aligned}$$

for $x = 1, 2, 3, \ldots$

Next

$$x = 1 \Longrightarrow f(1) = \frac{\beta}{1} \frac{\alpha}{\beta} f(0)$$

$$x = 2 \Longrightarrow f(2) = \frac{\beta}{2} [\frac{\alpha}{\beta} + 1] f(1) = \frac{\beta^2}{2!} (\frac{\alpha}{\beta} + 1)(\frac{\alpha}{\beta}) f(0)$$

$$x = 3 \Longrightarrow f(3) = \frac{\beta}{3} [\frac{\alpha}{\beta} + 2] f(2) = \frac{\beta^3}{3!} (\frac{\alpha}{\beta} + 2)(\frac{\alpha}{\beta} + 1)(\frac{\alpha}{\beta}) f(0)$$

In general,

$$f(x) = \frac{\beta^x}{x!}(\frac{\alpha}{\beta}+x-1)(\frac{\alpha}{\beta}+x-2)\dots(\frac{\alpha}{\beta}+2)(\frac{\alpha}{\beta}+1)\frac{\alpha}{\beta}f(0)$$
$$= \frac{\beta^x}{x!}(\frac{\alpha}{\beta}+x-1)(\frac{\alpha}{\beta}+x-2)\dots(\frac{\alpha}{\beta}+2)(\frac{\alpha}{\beta}+1)\frac{\alpha}{\beta}\frac{\Gamma\frac{\alpha}{\beta}}{\Gamma\frac{\alpha}{\beta}}f(0)$$

Therefore

$$f(x) = \frac{\beta^x}{x!}(\frac{\alpha}{\beta}+x-1)(\frac{\alpha}{\beta}+x-2)\dots(\frac{\alpha}{\beta}+2)(\frac{\alpha}{\beta}+1)\frac{\Gamma(\frac{\alpha}{\beta}+1}{)}\Gamma(\frac{\alpha}{\beta})f(0)$$
$$= \frac{\beta^x}{x!\Gamma\frac{\alpha}{\beta}}(\frac{\alpha}{\beta}+x-1)(\frac{\alpha}{\beta}+x-2)\dots(\frac{\alpha}{\beta}+2)(\frac{\alpha}{\beta}+1)\Gamma(\frac{\alpha}{\beta}+2)f(0)$$
$$= \vdots$$
$$= \frac{\beta^x}{x!\Gamma\frac{\alpha}{\beta}}(\frac{\alpha}{\beta}+x-1)\Gamma(\frac{\alpha}{\beta}+x-1)f(0)$$
$$= \frac{\beta^x}{x!\Gamma\frac{\alpha}{\beta}}\Gamma(\frac{\alpha}{\beta}+x)f(0); x = 1,2,3,\dots$$

$\Longrightarrow$

$$f(x) = \frac{\beta^x}{x!\Gamma\frac{\alpha}{\beta}}\Gamma(\frac{\alpha}{\beta}+x)f(0); x = 1,2,3,\dots \quad (4.18)$$

$$\Longrightarrow \sum_{x=0}^{\infty} f(x) = \frac{f(0)}{\Gamma\frac{\alpha}{\beta}}\sum_{x=0}^{\infty}\frac{\beta^x}{x!}\Gamma(\frac{\alpha}{\beta}+x)$$
$$\Longrightarrow 1 = \frac{f(0)}{\Gamma\frac{\alpha}{\beta}}\sum_{x=0}^{\infty}\frac{\beta^x}{x!}\Gamma(\frac{\alpha}{\beta}+x)$$

$\Longrightarrow$

$$\Gamma\frac{\alpha}{\beta} = f(0)[\Gamma\frac{\alpha}{\beta} + \frac{\beta}{1}\Gamma(\frac{\alpha}{\beta}+1) + \frac{\beta^2}{2!}\Gamma(\frac{\alpha}{\beta}+2) + \frac{\beta^3}{3!}\Gamma(\frac{\alpha}{\beta}+3) + \dots]$$
$$= f(0)[\Gamma\frac{\alpha}{\beta} + \frac{\beta}{1}\frac{\alpha}{\beta}\Gamma(\frac{\alpha}{\beta}) + \frac{\beta^2}{2!}(\frac{\alpha}{\beta}+1)\frac{\alpha}{\beta}\Gamma(\frac{\alpha}{\beta}) + \frac{\beta^3}{3!}(\frac{\alpha}{\beta}+2)(\frac{\alpha}{\beta}+1)\frac{\alpha}{\beta}\Gamma(\frac{\alpha}{\beta}) + \dots]$$
$$= \frac{\alpha}{\beta}f(0)[1 + \beta\frac{\alpha}{\beta} + \frac{\beta^2}{2!}(\frac{\alpha}{\beta}+1)\frac{\alpha}{\beta} + \frac{\beta^3}{3!}(\frac{\alpha}{\beta}+2)(\frac{\alpha}{\beta}+1)\frac{\alpha}{\beta} + \dots]$$

Therefore

$$1 = f(0)[1 + \beta\binom{\frac{\alpha}{\beta}}{1} + \beta^2\binom{\frac{\alpha}{\beta}+1}{2} + \beta^3\binom{\frac{\alpha}{\beta}+2}{3} + \ldots]$$
$$= f(0)\sum_{x=0}^{\infty}\beta^x\binom{\frac{\alpha}{\beta}+x-1}{x}$$

Thus

$$1 = f(0)\sum_{x=0}^{\infty}\beta^x[(-1)^x\binom{-\frac{\alpha}{\beta}}{x}]$$

[Using the identity $\binom{r+k-1}{k} = (-1)^k\binom{-r}{k}$]
Therefore

$$1 = f(0)\sum_{x=0}^{\infty}(-\beta)^x\binom{-\frac{\alpha}{\beta}}{x} = f(0)(1-\beta)^{-\frac{\alpha}{\beta}}$$

Hence

$$f(0) \quad = \quad (1-\beta)^{\frac{\alpha}{\beta}} \tag{4.19}$$

Therefore

$$f(x) \quad = \quad \frac{\beta^x}{x!\Gamma\frac{\alpha}{\beta}}\Gamma(\frac{\alpha}{\beta}+x)(1-\beta)^{\frac{\alpha}{\beta}}; x = 1, 2, 3, \ldots \tag{4.20}$$

There are now two sub-cases to be considered.

Sub-Case 1: When $0 < \beta < 1$

Let $\frac{\alpha}{\beta} = r$ and $\beta = p$.

Then equation(10) becomes

$$f(x) = \frac{p^x}{x!}\frac{\Gamma(r+x)}{\Gamma r}(1-p)^r$$
$$= \frac{\Gamma(r+x)}{x!\Gamma r}p^x(1-p)^r$$
$$= \frac{(r+x-1)(r+x-2)\ldots(r+1)r\Gamma r}{12\ldots x\Gamma r}p^x(1-p)^r$$
$$= \binom{r+x-1}{x}p^x(1-p)^r; 0 < p < 1; x = 0, 1, 2, \ldots$$

i.e

$$f(x) \quad = \quad \binom{r+x-1}{x}p^x(1-p)^r; 0 < p < 1; x = 0, 1, 2, \ldots \tag{4.21}$$

which is the Negative Binomial distribution.

If r is a positive integer then equation(11) gives the probability that x successes precede the r^{th} failure in an infinite sequence of Bernoulli trials with probability of success in each trial equal to p.

Sub-Case 2: When $\beta < 0$

Let $\frac{\alpha}{\beta} = -\nu$ and $\beta = \frac{p}{p-1}$ where $\nu > 0$ and $0 < p < 1$.

Then equation(10) becomes

$$\begin{aligned}
f(x) &= \frac{(\frac{p}{p-1})^x}{x!}\frac{\Gamma(-\nu+x)}{\Gamma-\nu}(1-\frac{p}{p-1})^{-\nu} \\
&= \frac{\Gamma(-\nu+x)}{x!\Gamma-\nu}(\frac{p}{p-1})^x(1-\frac{p}{p-1})^{-\nu} \\
&= \frac{\Gamma(-\nu+x)}{x!\Gamma-\nu}(\frac{p}{p-1})^x(1-\frac{p-1-p}{p-1})^{-\nu} \\
&= \frac{\Gamma(-\nu+x)}{x!\Gamma-\nu}p^x(\frac{1}{p-1})^x(\frac{-1}{p-1})^{-\nu} \\
&= \frac{\Gamma(-\nu+x)}{x!\Gamma-\nu}(-1)^{-\nu}p^x(\frac{1}{p-1})^{-\nu+x} \\
&= \frac{\Gamma(-\nu+x)}{x!\Gamma-\nu}(-1)^{-\nu}p^x(\frac{-1}{1-p})^{-\nu+x} \\
&= \frac{\Gamma(-\nu+x)}{x!\Gamma-\nu}(-1)^{-\nu}(-1)^{-\nu+x}p^x(1-p)^{\nu-x} \\
&= \frac{\Gamma(-\nu+x)}{x!\Gamma-\nu}(-1)^{-2\nu+x}p^x(1-p)^{\nu-x} \\
&= \frac{\Gamma(-\nu+x)}{x!\Gamma-\nu}(-1)^x p^x(1-p)^{\nu-x}
\end{aligned}$$

Therefore

$$\begin{aligned}
f(x) &= (-1)^x\frac{(-\nu+x-1)(-\nu+x-2)\dots(-\nu+x-x)\Gamma(-\nu+x-x)}{12\dots x\Gamma-\nu}p^x(1-p)^{\nu-x} \\
&= (-1)^x\frac{(-\nu+x-1)(-\nu+x-2)\dots(-\nu)\Gamma(-\nu)}{12\dots x\Gamma-\nu}p^x(1-p)^{\nu-x} \\
&= (-1)^x\binom{-\nu+x-1}{x}p^x(1-p)^{\nu-x} \\
&= \binom{\nu}{x}p^x(1-p)^{\nu-x}
\end{aligned}$$

[Using the identity $\binom{-r}{k} = (-1)^k\binom{r+k-1}{k}$]

Hence

$$f(x) = \binom{\nu}{x} p^x (1-p)^{\nu - x} \quad for \; x = 0, 1, 2, \ldots, \nu \tag{4.22}$$

which is a binomial distribution with parameters ν and p.

We can summarize the above discussion as

<u>Theorem 1</u>

Assume that

$$p_n = \frac{a + bn}{n} p_{n-1} \quad ; n = 1, 2, 3, \ldots$$

holds.Then we have one of the three cases.

$$p_n = \frac{e^{-\alpha}\alpha^n}{n!} \qquad (\alpha > 0)$$

$$p_n = \binom{r+n-1}{n} p^n (1-p)^r \qquad 0 < p < 1$$

$$p_n = \binom{\nu}{n} p^n (1-p)^{\nu - n} \qquad 0 < p < 1 \; and \; \nu > 0$$

4.3.2 Willmot Pattern

Let

$$p_n = (a + \frac{b}{n}) p_{n-1} \quad ; n = 2, 3, \ldots$$

and $p_1 > 0, p_0 = 0 \Longrightarrow$

$$\begin{aligned} np_n &= (a + bn) p_{n-1} \quad ; n = 2, 3, \ldots \\ &= [a(n - 1 + 1) + b] p_{n-1} \\ &= a(n-1) p_{n-1} + b p_{n-1} \end{aligned}$$

Therefore

$$\sum_{n=2}^{\infty} np_n s^n = a \sum_{n=2}^{\infty} (n-1) p_{n-1} s^n + (a+b) \sum_{n=2}^{\infty} p_{n-1} s^n$$

$$s \sum_{n=2}^{\infty} np_n s^{n-1} = as^2 \sum_{n=2}^{\infty} (n-1) p_{n-1} s^{n-2} + (a+b) s \sum_{n=2}^{\infty} p_{n-1} s^{n-1}$$

$$\sum_{n=2}^{\infty} np_n s^{n-1} = as\sum_{n=2}^{\infty}(n-1)p_{n-1}s^{n-2} + (a+b)\sum_{n=2}^{\infty} p_{n-1}s^{n-1}$$

$$\frac{\partial G}{\partial s} - p_1 = as\frac{\partial G}{\partial s} + (a+b)[G(s) - p_0]$$

But $p_0 = 0$
Therefore

$$\frac{\partial G}{\partial s} - p_1 = as\frac{\partial G}{\partial s} + (a+b)G(s)$$

$$(1-as)\frac{\partial G}{\partial s} - (a+b)G(s) = p_1 \tag{4.23}$$

Solving this differential equation in pgf ,we consider the following cases:-

Case (i):When $a = 0$ and $b = 0$
Then (4.23) becomes

$$\frac{\partial G}{\partial s} = p_1 \Longrightarrow G(s) = p_1 s + c$$

Therefore

$$G(0) = c$$

But $G(0) = p_0 = 0 \Longrightarrow C = 0$

$$G(s) = p_1 s \Longrightarrow G(1) = p_1 = 1$$

$$G(s) = s$$

Thus

$$p_n = \begin{cases} 1 & when \ \ n=1 \\ 0 & when \ \ n\neq 1 \end{cases}$$

Case (ii):When $a = 0$ and $b \neq 0$
Then

$$\frac{\partial G}{\partial s} - bG(s) = p_1$$

$$I = IntegratingFactor = e^{-\int b ds} = e^{-bs}$$

Therefore

$$e^{-bs}\frac{\partial G}{\partial s} - be^{-bs}G(s) = p_1 e^{-bs}$$

i.e.

$$\frac{\partial}{\partial s}[e^{-bs}G(s)] = p_1 e^{-bs}$$

$$\Longrightarrow e^{-bs}G(s) = \int p_1 e^{-bs} ds = -\frac{p_1}{b}e^{-bs} + c$$

Therefore

$$G(0) = -\frac{p_1}{b} + C \Longrightarrow 0 = -\frac{p_1}{b} + C \Longrightarrow C = \frac{p_1}{b}$$

and

$$e^{-b}G(1) = -\frac{p_1}{b}e^{-b} + \frac{p_1}{b}$$

i.e.

$$1 = \frac{p_1}{b}(e^b - 1)$$

Hence

$$p_1 = \frac{b}{e^b - 1}$$

Therefore

$$\begin{aligned} e^{-bs}G(s) &= c - \frac{p_1}{b}e^{-bs} \\ &= \frac{p_1}{b} - \frac{p_1}{b}e^{-bs} \\ &= \frac{p_1}{b}[1 - e^{-bs}] \end{aligned}$$

Thus

$$\begin{aligned} G(s) &= \frac{p_1}{b}[e^{bs} - 1] \\ &= \frac{e^{bs} - 1}{e^b - 1} \\ &= \frac{1}{e^b - 1}[\sum_{n=0}^{\infty} \frac{(bs)^n}{n!} - 1] \end{aligned}$$

Therefore

$$G(s) = \frac{1}{e^b - 1}[\sum_{n=1}^{\infty} \frac{b^n s^n}{n!}$$

Hence

$$p_n = \frac{b^n}{n!(e^b - 1)} \;; n = 1, 2, 3, \ldots \;\; p_0 = 0$$

which is the Zero-truncated Poisson distribution with parameter b.

Case (iii): When $a \neq 0$ and $b = 0$

Then

$$(1-as)\frac{\partial G}{\partial s} - aG(s) = p_1$$

$$\Longrightarrow \frac{\partial}{\partial s}[(1-as)G(s)] = p_1 \Longrightarrow (1-as)G(s) = p_1 + C$$

$$G(0) = C, i.e. 0 = C$$

Therefore

$$G(s) = \frac{p_1 s}{(1-as)} \Longrightarrow 1 = G(1) = \frac{p_1}{(1-a)} \Longrightarrow p_1 = 1-a$$

Hence

$$G(s) = \frac{(1-a)s}{(1-as)}$$

which implies that

$$p_n = a^{n-1}(1-a) \; ; n = 1, 2, \ldots \; p_0 = 0$$

Case (iv): $a \neq 0$ and $b = 0$ and $(a+b) \geq 0$

Then

$$(1-as)\frac{\partial G}{\partial s} - (a+b)G(s) = p_1$$

When $a + b = 0$, then $a \neq 0$

$$(1-as)\frac{\partial G}{\partial s} = p_1$$

Therefore

$$\frac{\partial G}{\partial s} = \frac{p_1}{1-as}$$

which implies that

$$\begin{aligned} G(s) &= \int \frac{p_1}{1-as} ds + C \\ &= \frac{1}{a}\int \frac{-a}{1-as} ds + C \\ &= \frac{p_1}{a} log(1-as) + C \end{aligned}$$

$$1 = G(1) = \frac{p_1}{a} log(1-a) + C = \frac{p_1}{a} C_1 log(1-as)$$

$$\Longrightarrow C_1 = \frac{a}{p_1 log(1-a)}$$

Thus

$$G(s) = \frac{log(1-as)}{log(1-a)}$$

which is the pgf of a logarithmic distribution with probability a i.e.

$$p_n = \frac{a^n}{-nlog(1-a)} \ ; n = 1, 2, 3, \ldots, 0 < a < 1$$

When $a + b > 0$,**then** $a \neq 0$

$$\frac{\partial G}{\partial s} - \frac{a+b}{(1-as)} G(s) = \frac{p_1}{(1-as)}$$

Therefore

$$I = e^{-\int(\frac{a+b}{(1-as)})ds} = e^{\frac{(a+b)}{a}\int -\frac{a}{1-as}ds} = e^{\frac{(a+b)}{a} log(1-as)}$$

i.e.

$$I = (1-as)^{\frac{(a+b)}{a}} = (1-as)^m$$

where $m = \frac{(a+b)}{a}$
Therefore

$$(1-as)^m \frac{\partial G}{\partial s} - ma(1-as)^{m-1} G(s) = p_1 (1-as)^{m-1}$$

$$\frac{\partial}{\partial s}(1-as)^m G(s) = p_1 (1-as)^{m-1}$$

$$\Longrightarrow (1-as)^m G(s) = p_1 \int (1-as)^{m-1} ds + C$$

Let

$$u = (1-as) \Longrightarrow du = -ads$$

Hence

$$u^m G(s) = p_1 \int \frac{u^{m-1}}{-a} du + C$$

$$= -\frac{p_1}{a}\frac{u^m}{m} + C$$

$$s = 0 \Longrightarrow u = 1$$

Therefore

$$0 = G(0) = -\frac{p_1}{a}\frac{1}{m} + C$$

$$\Longrightarrow C = \frac{p_1}{ma}$$

$$u^m G(s) = -\frac{p_1}{a}\frac{u^m}{m} + \frac{p_1}{ma}$$

i.e.

$$\begin{aligned} G(s) &= \frac{p_1}{ma}[1 - u^m] \\ &= \frac{p_1}{ma}[u^{-m} - 1] \\ &= \frac{p_1}{ma}[(1 - as)^{-m} - 1] \end{aligned}$$

Therefore

$$1 = G(1) = \frac{p_1}{ma}[(1 - a)^{-m} - 1]$$

$$\Longrightarrow ma = p_1[(1 - a)^{-m} - 1]$$

$$\Longrightarrow p_1 = \frac{ma}{(1 - a)^{-m} - 1}$$

Hence

$$G(s) = \frac{1}{(1 - a)^{-m} - 1}[(1 - as)^{-m} - 1]$$

OR

$$G(s) = \frac{1 - (1 - as)^{-m}}{1 - (1 - a)^{-m}}$$

Note that m can be
a positive integer
a negative integer
a fraction

When m is a positive integer,then

$$G(s) = \frac{1 - \sum_{n=0}^{\infty} \binom{-m}{n}(-as)^n}{1 - \sum_{n=0}^{\infty} \binom{-m}{n}(-a)^n}$$
$$= \frac{\sum_{n=1}^{\infty} \binom{-m}{n}(-as)^n}{\sum_{n=1}^{\infty} \binom{-m}{n}(-a)^n}$$
$$= \frac{\sum_{n=1}^{\infty} (-1)^n \binom{-m}{n} a^n s^n}{\sum_{n=1}^{\infty} (-1)^n \binom{-m}{n} a^n}$$
$$= \frac{\sum_{n=1}^{\infty} \binom{m+n-1}{n} a^n s^n}{\sum_{n=1}^{\infty} \binom{m+n-1}{n} a^n}$$

Therefore

$$p_n = \frac{\binom{m+n-1}{n} a^n}{\sum_{n=1}^{\infty} \binom{m+n-1}{n} a^n} \ ; n = 1, 2, \ldots \quad p_0 = 0$$

When m is a negative integer,let r=-mwhich is a positive integer. Therefore

$$G(s) = \frac{1 - (1 - as)^r}{1 - (1 - a)^r} = \frac{\sum_{n=1}^{r} \binom{r}{n}(-as)^n}{\sum_{k=1}^{r} \binom{r}{k}(-a)^k}$$

implying that

$$p_n = \frac{(-a)^n \binom{r}{n}}{\sum_{k=1}^{r} \binom{r}{k}(-a)^k} \ ; n = 1, 2, \ldots, r \quad a < 0$$

We can summarize the above discussion as

<u>Theorem</u>

Let

$$p_n = (a + \frac{b}{n}) p_{n-1} \ ; n = 2.3. \ldots \quad p_1 > 0$$

Then

a)

$$(1 - as)\frac{\partial G}{\partial s} - (a + b)G(s) = p_1$$

and

b)

(i)

$$p_n = \begin{cases} 1 & when \;\; n=1 \\ 0 & when \;\; n\neq 1 \end{cases} \qquad a=0,b=0$$

(ii)

$$p_n = \begin{cases} \frac{b^n}{n!(e^b-1)} & ;n=1,2,3,\ldots \\ 0 & n=0 \end{cases} \qquad a=0,b\neq 0$$

(iii)

$$p_n = \begin{cases} a^{n-1}(1-a) & ;n=1,2,\ldots \\ 0 & n=0 \end{cases} \qquad a\neq 0,b=0$$

(iv)

$$p_n = \begin{cases} \frac{a^n}{-n\log(1-a)} & ;n=1,2,3,\ldots \\ 0 & n=0 \end{cases} \qquad a\neq 0,b\neq 0$$

(v)

$$p_n = \begin{cases} \frac{\binom{m+n-1}{n}a^n}{\sum_{n=1}^{\infty}\binom{m+n-1}{n}a^n} & ;n=1,2,\ldots \qquad m=\frac{(a+b)}{a}\in\aleph,a\neq 0,b\neq 0,a+b>0 \\ \frac{(-a)^n\binom{r}{n}}{\sum_{k=1}^{r}\binom{r}{k}(-a)^k} & ;n=1,2,\ldots,r \qquad m=-r,r\in\aleph,a\neq 0,b\neq 0,a+b>0 \end{cases}$$

4.3.3 Panjer Distribution of order k

Let

$$q_{n+1} = (a+\frac{b}{n+1})q_n \;\; for \;\; n\geq k \tag{4.24}$$

where

$$q_n = 0 \;\; for \;\; n\leq k-1 \tag{4.25}$$

This pattern is said to be Panjer distribution with parameters a,b $\in \Re$ and k $\in \aleph$;or simply Panjer distribution of order k.

We now wish to characterize this distribution by a differential equation for its probability generating function.
This result will be used to to identify all distributions of the Panjer class of order k.
Define

$$G(s) = \sum_{n=0}^{\infty} q_n s^n \tag{4.26}$$

Then

$$q_n = \frac{G^{(n)}(0)}{n!} \qquad (4.27)$$

where

$$G^{(n)}(0) = \frac{d^nG}{ds^n}\Big|_{s=0} \qquad (4.28)$$

From (4.26)

$$\frac{dG}{ds} = \sum_{n=0}^{\infty} nq_n s^{n-1} = \sum_{n=1}^{\infty} nq_n s^{n-1}$$
$$\frac{d^2G}{ds^2} = \sum_{n=1}^{\infty}(n-1)nq_n s^{n-2} = \sum_{n=2}^{\infty}(n-1)nq_n s^{n-2}$$
$$\frac{d^3G}{ds^3} = \sum_{n=2}^{\infty}(n-2)(n-1)nq_n s^{n-3} = \sum_{n=3}^{\infty}(n-2)(n-1)nq_n s^{n-3}$$

Therefore

$$\frac{d^lG}{ds^l} = \sum_{n=l}^{\infty}(n-l+1)\dots(n-1)nq_n s^{n-l}$$

$$\frac{1}{l!}\frac{d^lG}{ds^l} = \sum_{n=l}^{\infty} \frac{(n-l+1)(n-l+2)\dots(n-1)nq_n s^{n-l}}{1 \quad 2 \quad \dots \quad l}$$
$$= \sum_{n=l}^{\infty} \binom{n}{l} q_n s^{n-l}$$

But $q_n = 0 \ \ for \ \ n \leq k-1$
Therefore

$$\frac{1}{l!}\frac{d^lG}{ds^l} = \sum_{n=k}^{\infty} \binom{n}{l} q_n s^{n-l} \qquad (4.29a)$$
$$= \binom{k}{l} q_k s^{k-l} + \sum_{n=k+1}^{\infty} \binom{n}{l} q_n s^{n-l}$$

Thus

$$\frac{1}{l!}\frac{d^lG}{ds^l} - \binom{k}{l}q_k s^{k-l} = \sum_{n=k+1}^{\infty} \binom{n}{l} q_n s^{n-l} \qquad (4.29b)$$

Equation(4.24)implies that

$$q_n = (a + \frac{b}{n})q_{n-1} \ \ for \ \ n-1 \geq k$$

Substitute this in (4.29b)

$$\frac{1}{l!}\frac{d^lG}{ds^l} - \binom{k}{l}q_k s^{k-l} = \sum_{n=k+1}^{\infty} \binom{n}{l}(a + \frac{b}{n})q_{n-1}s^{n-l}$$

Put $n = j + 1 \Longrightarrow j = n - 1$
Therefore

$$\frac{1}{l!}\frac{d^lG}{ds^l} - \binom{k}{l}q_k s^{k-l} = \sum_{j=k}^{\infty} \binom{j+1}{l}(a + \frac{b}{j+1})q_j s^{j+1-l}$$

$$= \sum_{j=k}^{\infty} a q_j \binom{j+1}{l} s^{j+1-l} + \sum_{j=k}^{\infty} \frac{b}{j+1} q_j \binom{j+1}{l} s^{j+1-l}$$

$$\frac{1}{l!}\frac{d^lG}{ds^l} - q_k\binom{k}{l}s^{k-l} = as\sum_{j=k}^{\infty} q_j\binom{j+1}{l}s^{j-l} + b\sum_{j=k}^{\infty}\frac{q_j}{j+1}\binom{j+1}{l}s^{j+1-l}$$

$$= as[\sum_{j=k}^{\infty} q_j\frac{(j+1)j!}{l!(j+1-l)!}s^{j-l}] + b\sum_{j=k}^{\infty}[\frac{q_j}{j+1}\frac{(j+1)j!}{l!(j+1-l)!}s^{j+1-l}]$$

$$= as[\sum_{j=k}^{\infty} q_j(j+1)\frac{j!}{l!(j+1-l)(j-l)!}s^{j-l}]$$

$$+ b\sum_{j=k}^{\infty}[q_j\frac{j!}{l(l-1)![j-(l-1)]!}s^{j-(l-1)}]$$

$$= as\sum_{j=k}^{\infty}[q_j\frac{(j+1)}{(j-l+1)}\binom{j}{l}s^{j-l}] + b\sum_{j=k}^{\infty}[\frac{q_j}{l}\binom{j}{l-1}s^{j-(l-1)}]$$

$$= as\sum_{j=k}^{\infty}[q_j\frac{(j-l+1+l)}{(j-l+1)}\binom{j}{l}s^{j-l}] + \frac{b}{l}\sum_{j=k}^{\infty}[q_j\binom{j}{l-1}s^{j-(l-1)}]$$

$$= as\sum_{j=k}^{\infty}[q_j\binom{j}{l}s^{j-l} + q_j\frac{l}{(j-l+1)}\binom{j}{l}s^{j-l}] + \frac{b}{l}\sum_{j=k}^{\infty}[q_j\binom{j}{l-1}s^{j-(l-1)}]$$

$$= as\sum_{j=k}^{\infty} q_j\binom{j}{l}s^{j-l} + a\sum_{j=k}^{\infty} q_j\frac{l}{(j-l+1)}\binom{j}{l}s^{j-l+1} + \frac{b}{l}\sum_{j=k}^{\infty} q_j\binom{j}{l-1}s^{j-}$$

$$\frac{1}{l!}\frac{d^lG}{ds^l} - q_k\binom{k}{l}s^{k-l} = as\sum_{j=k}^{\infty} q_j\binom{j}{l}s^{j-l} + a\sum_{j=k}^{\infty} q_j\frac{l}{(j-l+1)}\frac{j!}{l(l-1)!(j-l)!}s^{j-l+1}$$
$$+\frac{b}{l}\sum_{j=k}^{\infty} q_j\binom{j}{l-1}s^{j-(l-1)}$$
$$= as\sum_{j=k}^{\infty} q_j\binom{j}{l}s^{j-l} + a\sum_{j=k}^{\infty} q_j\frac{j!}{(j-l+1)(l-1)!(j-l)!}s^{j-l+1}$$
$$+\frac{b}{l}\sum_{j=k}^{\infty} q_j\binom{j}{l-1}s^{j-(l-1)}$$
$$= as\sum_{j=k}^{\infty} q_j\binom{j}{l}s^{j-l} + a\sum_{j=k}^{\infty} q_j\frac{j!}{(l-1)!(j-l+1)!}s^{j-l+1} + \frac{b}{l}\sum_{j=k}^{\infty} q_j\binom{j}{l-1}s^{j-(l-1)}$$
$$= as\sum_{j=k}^{\infty} q_j\binom{j}{l}s^{j-l} + a\sum_{j=k}^{\infty} q_j\binom{j}{(l-1)}s^{j-(l-1)} + \frac{b}{l}\sum_{j=k}^{\infty} q_j\binom{j}{l-1}s^{j-(l-1)}$$
$$= as\sum_{j=k}^{\infty} q_j\binom{j}{l}s^{j-l} + (a+\frac{b}{l})\sum_{j=k}^{\infty} q_j\binom{j}{(l-1)}s^{j-(l-1)}$$

Thus

$$\frac{1}{l!}\frac{d^lG}{ds^l} - q_k\binom{k}{l}s^{k-l} = as\sum_{j=k}^{\infty} q_j\binom{j}{l}s^{j-l} + (a+\frac{b}{l})\sum_{j=k}^{\infty} q_j\binom{j}{(l-1)}s^{j-(l-1)} \quad (4.29)$$

Using (4.29a),then (4.29b) becomes

$$\frac{1}{l!}\frac{d^lG}{ds^l} - q_k\binom{k}{l}s^{k-l} = \frac{as}{l!}\frac{d^lG}{ds^l} + (a+\frac{b}{l})\frac{1}{(l-1)!}\frac{d^{l-1}G}{ds^{l-1}}$$

Therefore

$$\frac{1-as}{l!}\frac{d^lG}{ds^l} = \frac{(a+\frac{b}{l})}{(l-1)!}\frac{d^{l-1}G}{ds^{l-1}} + q_k\binom{k}{l}s^{k-l}$$

Thus

$$(1-as)G^{(l)}(s) = (la+b)G^{(l-1)}(s) + q_k\binom{k}{l}s^{k-l} \quad (4.30)$$

Put $l = k+1$ **equation (4.30) becomes**

$$(1-as)G^{(k+1)}(s) = [(k+1)a+b]G^{(k)}(s) \quad (4.31)$$

Putting $s = 0$ then (4.31) becomes

$$G^{(k+1)}(0) = [(k+1)a + b]G^{(k)}(0)$$

Assuming that $k = n$

$$\Longrightarrow G^{(n+1)}(0) = [(n+1)a + b]G^{(n)}(0)$$

Therefore

$$\frac{G^{(n+1)}(0)}{(n+1)!} = [(n+1)a + b]\frac{G^{(n)}(0)}{(n+1)!}$$

$$\Longrightarrow q_{n+1} = [\frac{(n+1)a + b}{(n+1)!}]\frac{G^{(n)}(0)}{(n)!} = [a + \frac{b}{n+1}]q_n$$

$$q_{n+1} = [a + \frac{b}{n+1}]q_n \qquad (4.32)$$

For equation (4.31),when $k = 0$ we get

$$(1 - as)G^1(s) = [a + b]G(s)$$

i.e.

$$(1 - as)\frac{\partial G(s)}{\partial s} - (a + b)G(s) = 0$$

which is the differential equation for the original Panjer's pattern.

When $k = 1$ equation(4.31) becomes

$$(1 - as)\frac{\partial^2 G(s)}{\partial s^2} = (2a + b)\frac{\partial G(s)}{\partial s}$$

Put

$$\frac{\partial}{\partial s}[(1 - as)\frac{\partial G(s)}{\partial s}] = (1 - as)\frac{\partial^2 G(s)}{\partial s^2} - a\frac{\partial G(s)}{\partial s}$$

Therefore

$$\frac{\partial}{\partial s}[(1 - as)\frac{\partial G(s)}{\partial s}] + a\frac{\partial G(s)}{\partial s} = (1 - as)\frac{\partial^2 G(s)}{\partial s^2} = (2a + b)\frac{\partial G(s)}{\partial s}$$

$$\Longrightarrow \frac{\partial}{\partial s}[(1 - as)\frac{\partial G(s)}{\partial s}] - (a + b)\frac{\partial G(s)}{\partial s} = 0$$

$$\frac{\partial}{\partial s}[(1-as)\frac{\partial G(s)}{\partial s}-(a+b)G(s)]=0$$

$$(1-as)\frac{\partial G(s)}{\partial s}-(a+b)G(s)=c$$

For $s=0$,we have

$$(1-a)\frac{\partial G(s)}{\partial s}\mid_{s=0}-(a+b)G(0)=c$$

But

$$p_k=\frac{G^{(k)}(0)}{k!}\Longrightarrow k!p_k=G^{(k)}(0)$$

$$\Longrightarrow p_1=G^1(0)=\frac{\partial G(s)}{\partial s}\mid_{s=0}$$

and $p_0=G(0)$
But $p_n=0$ for $n\leq k-1\Longrightarrow n\leq 0$
Therefore $p_0=0$ and $p_1=c$ Thus

$$(1-as)\frac{\partial G(s)}{\partial s}-(a+b)G(s)=p_1$$

which is equation (4.23).

4.3.4 Compound Distributions

In Chapter 2 we have discussed compound distributions based on the sums of a random number of iid random variables.We let

$$S_N=X_1+X_2+\ldots+X_N$$

where the $X_i's$ are independent and identically distributed (iid) random variables and N is also a random variable independent of the $X_i's$.

If G(s),F(s) and H(s) are the pgf's of X,N and S_N respectively,then

$$H(s)=F_N[G_X(s)]$$

which is a compound distribution.

Further if X_i is Bernoulli with parameter π,then

$$G(s)=G_X(s)=(1-\pi+\pi s)$$

Therefore

$$H(t) = F(1 - \pi + \pi s)$$

For Compound Poisson Distribution,let N be Poisson with parameter λ.Then

$$\begin{aligned} H(s) &= e^{-\lambda\pi(1-s)} \\ &\equiv F[\beta(1-s)] \end{aligned}$$

where $\beta = \lambda\pi$

For Compound Binomial Distribution,let N be Binomial with parameters m and p.Then

$$\begin{aligned} H(s) &= [1 - \pi p(1-s)]^m \\ &= F[\beta(1-s)] \end{aligned}$$

where $\beta = \pi p$

For Compound Negative Binomial Distribution,let N be Negative Binomial with parameters p and r.Then

$$\begin{aligned} H(s) &= [\frac{1}{1 - \beta(1-s)}]^r \\ &= F[\beta(1-s)] \end{aligned}$$

where $\beta = -\frac{(1-p)\pi}{p}$

Thus compound distributions of the form

$$H(s) = F[\beta(1-s)]$$

come from Poisson,Binomial and Negative Binomial Distributions.

Chapter 5

MAXIMUM LIKELIHOOD ESTIMATION

5.1 Sprott Approach

Sprott(1958,1965,1983) developed a method for estimating parameters in recursive relations and difference-differential equations.

We wish to review some of these papers to grasp the concept.

5.1.1 Estimating parameters of a convolution by Maximum Likelihood

Let

$$Z = X + Y$$

and

$$
\begin{aligned}
p_k &= Prob(Z = k) \\
&= Prob(X + Y = k) \\
&= \sum_{i=0}^{k} Prob(X = i, Y = k - i) \\
&= \sum_{i=0}^{k} Prob(X = i) Prob(Y = k - i) \\
&= \sum_{i=0}^{k} r_i q_{k-i} \\
&= \sum_{i=0}^{k} \binom{N}{i} \phi^i (1-\phi)^{N-i} \frac{e^{-\theta}\theta^{k-i}}{(k-i)!} \\
&= e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \phi^i (1-\phi)^{N-i}
\end{aligned}
$$

Thus

$$
p_k = e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \phi^i (1-\phi)^{N-i} \tag{5.1}
$$

Therefore

$$
\begin{aligned}
\frac{\partial p_k}{\partial \theta} &= -e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \phi^i (1-\phi)^{N-i} + e^{-\theta} \sum_{i=0}^{k} \frac{(k-i)\theta^{k-i-1}}{(k-i)!} \binom{N}{i} \phi^i (1-\phi)^{N-i} \\
&= -p_k + e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i-1}}{(k-i-1)!} \binom{N}{i} \phi^i (1-\phi)^{N-i}
\end{aligned}
$$

Thus

$$
\frac{\partial p_k}{\partial \theta} = -p_k + p_{k-1} \tag{5.2}
$$

Therefore

$$
\frac{1}{p_k} \frac{\partial p_k}{\partial \theta} = -1 + \frac{p_{k-1}}{p_k}
$$

i.e.

$$\frac{\partial}{\partial\theta}log p_k = \frac{p_{k-1}}{p_k} - 1 \quad (5.3a)$$

$$= H_1(k) - 1 \quad (5.3b)$$

where

$$H_1 = \frac{p_{k-1}}{p_k}$$

Next

$$\frac{\partial p_k}{\partial \phi} = e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \frac{\partial}{\partial \phi} \phi^i (1-\phi)^{N-i}$$

$$= e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} [i\phi^{i-1}(1-\phi)^{N-i} - (N-i)\phi^i (1-\phi)^{N-i-1}]$$

$$= e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \phi^{i-1}(1-\phi)^{N-i-1}[i(1-\phi) - (N-i)\phi]$$

$$= e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \phi^{i-1}(1-\phi)^{N-i-1}[i - N\phi]$$

$$= e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} i \binom{N}{i} \phi^{i-1}(1-\phi)^{N-i-1} - N\phi e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \phi^i (1-\phi)^{N-i-1}$$

$$= e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} (i+k-k) \binom{N}{i} \frac{\phi^i (1-\phi)^{N-i}}{\phi(1-\phi)} - N\phi e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \frac{\phi^i (1-\phi)^{N-i}}{\phi(1-\phi)}$$

$$= \frac{e^{-\theta}}{\phi(1-\phi)} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} [-k+i+k] \binom{N}{i} \phi^i (1-\phi)^{N-i}$$

$$- \frac{N\phi}{\phi(1-\phi)} e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \phi^i (1-\phi)^{N-i}$$

$$= \frac{1}{\phi(1-\phi)} [-e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} (k-i) \binom{N}{i} \phi^i (1-\phi)^{N-i} + k e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i)!} \binom{N}{i} \phi^i (1-\phi)^{N-}$$

$$- \frac{N\phi}{\phi(1-\phi)} p_k$$

$$= \frac{1}{\phi(1-\phi)} [-e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i}}{(k-i-1)!} \binom{N}{i} \phi^i (1-\phi)^{N-i} + kp_k - N\phi p_k]$$

Therefore

$$\frac{\partial p_k}{\partial \phi} = \frac{1}{\phi(1-\phi)} [-\theta e^{-\theta} \sum_{i=0}^{k} \frac{\theta^{k-i-1}}{(k-i-1)!} \binom{N}{i} \phi^i (1-\phi)^{N-i} + kp_k - N\phi p_k]$$

$$= \frac{1}{\phi(1-\phi)} [-\theta p_{k-1} + kp_k - N\phi p_k]$$

Thus

$$\frac{\partial p_k}{\partial \phi} = \frac{1}{\phi(1-\phi)}[(k - N\phi)p_k - \theta p_{k-1}] \tag{5.3}$$

This implies that

$$\begin{aligned}\frac{1}{p_k}\frac{\partial p_k}{\partial \phi} &= \frac{1}{\phi(1-\phi)}[(k - N\phi) - \theta \frac{p_{k-1}}{p_k}] \\ &= \frac{1}{\phi(1-\phi)}[(k - N\phi) - \theta H_1(k)]\end{aligned}$$

i.e.

$$\frac{\partial}{\partial \phi} log p_k = \frac{1}{\phi(1-\phi)}[k - N\phi - \theta H_1(k)] \tag{5.4}$$

Now multiply (5.3b) and (5.4) by frequency a_k and sum the result over k:

For (5.3b)

$$\begin{aligned}\sum_k a_k \frac{\partial}{\partial \theta} log p_k &= \sum_k a_k [H_1(k) - 1] \\ &= \sum_k a_k H_1(k) - \sum_k a_k \\ &= \sum_k a_k H_1(k) - n\end{aligned}$$

where

$$n = \sum_k a_k$$

Let

$$S_\theta = \sum_k a_k \frac{\partial}{\partial \theta} log p_k$$

Therefore

$$S_\theta = \sum_k a_k H_1(k) - n \tag{5.5}$$

For (5.4)

$$\sum_k a_k \frac{\partial}{\partial\phi} logp_k = \frac{1}{\phi(1-\phi)} \sum_k a_k[k - N\phi - \theta H_1(k)]$$

$$= \frac{1}{\phi(1-\phi)}[\sum_k ka_k - N\phi \sum_k a_k - \theta \sum_k a_k H_1(k)]$$

$$= \frac{1}{\phi(1-\phi)}[n\frac{\sum_k ka_k}{n} - Nn\phi - \theta \sum_k a_k H_1(k)]$$

Thus

$$S_\phi = \frac{1}{\phi(1-\phi)}[n\bar{k} - Nn\phi - \theta \sum_k a_k H_1(k)] \qquad (5.6)$$

where

$$S_\phi = \sum_k a_k \frac{\partial}{\partial\phi} logp_k$$

$$\bar{k} = \frac{\sum_k ka_k}{n} = E(k) = E(Z)$$

and

$$n = \sum_k a_k$$

From (5.5) we have

$$\sum_k a_k H_1(k) = S_\theta + n$$

Substitute this in (5.6)and we get

$$S_\phi = \frac{1}{\phi(1-\phi)}[n\bar{k} - Nn\phi - \theta(S_\theta + n)] \qquad (5.7)$$

The Maximum Likelihood equations are $S_\theta = 0$ and $S_\phi = 0$.
Thus (5.7) becomes

$$0 = \frac{1}{\phi(1-\phi)}[n\bar{k} - Nn\phi - n\theta]$$

$$\Longrightarrow 0 = \bar{k} - N\phi - \theta$$

Therefore

$$\bar{k} = N\phi + \theta \tag{5.8}$$

i.e.

$$E(Z) = N\phi + \theta$$

which is the first moment equation.

This can be used to eliminate one of the parameters,say ϕ.Then S_θ of (5.5) can be written as a function of θ only,and hence the ML estimate can be found by iterating on S_θ.

Applying this method to Panjer(1981) Recursive pattern i.e.

$$p_k = \frac{(ak+b)}{k} p_{k-1} \;\; for \;\; k = 1, 2, \ldots$$

We have

$$\begin{aligned}
\frac{\partial p_k}{\partial a} &= p_{k-1} \frac{\partial}{\partial a}(\frac{ak+b}{k}) + \frac{(ak+b)}{k} \frac{\partial}{\partial a} p_{k-1} \\
&= p_{k-1} + \frac{(ak+b)}{k} \frac{\partial}{\partial a} p_{k-1} \\
&= p_{k-1} + \frac{p_k}{p_k - 1} \frac{\partial}{\partial a} p_{k-1}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial p_k}{\partial b} &= p_{k-1} \frac{\partial}{\partial b}(\frac{ak+b}{k}) + \frac{(ak+b)}{k} \frac{\partial}{\partial b} p_{k-1} \\
&= \frac{1}{k} p_{k-1} + \frac{(ak+b)}{k} \frac{\partial}{\partial b} p_{k-1} \\
&= \frac{1}{k} p_{k-1} + \frac{p_k}{p_k - 1} \frac{\partial}{\partial b} p_{k-1}
\end{aligned}$$

Thus

$$\begin{aligned}
\frac{1}{p_k} \frac{\partial p_k}{\partial a} &= \frac{p_{k-1}}{p_k} + \frac{1}{p_{k-1}} \frac{\partial}{\partial a} p_{k-1} \\
\Longleftrightarrow \frac{\partial}{\partial a} log p_k &= \frac{p_{k-1}}{p_k} + \frac{\partial}{\partial a} log p_{k-1}
\end{aligned}$$

and

$$\frac{1}{p_k} \frac{\partial p_k}{\partial b} = \frac{1}{k} \frac{p_{k-1}}{p_k} + \frac{1}{p_{k-1}} \frac{\partial}{\partial b} p_{k-1}$$

$$\iff \frac{\partial}{\partial b} log p_k = \frac{1}{k}\frac{p_{k-1}}{p_k} + \frac{\partial}{\partial b} log p_{k-1}$$

Therefore

$$\frac{\partial}{\partial a} log p_k = \frac{p_{k-1}}{p_k} + \frac{\partial}{\partial a} log p_{k-1} \qquad (5.9)$$

and

$$\frac{\partial}{\partial b} log p_k = \frac{1}{k}\frac{p_{k-1}}{p_k} + \frac{\partial}{\partial b} log p_{k-1} \qquad (5.10)$$

Now multiply (5.9) and (5.10) by frequency f_k and sum the result over k:

For (5.9)

$$\sum_k f_k \frac{\partial}{\partial a} log p_k = \sum_k f_k \frac{p_{k-1}}{p_k} + \sum_k f_k \frac{\partial}{\partial a} log p_{k-1}$$

Let

$$S_a = \sum_k f_k \frac{\partial}{\partial a} log p_k$$

Then

$$S_a = \sum_k f_k \frac{p_{k-1}}{p_k} + \sum_k f_k \frac{\partial}{\partial a} log p_{k-1} \qquad (5.11)$$

For (5.10)

$$\sum_k f_k \frac{\partial}{\partial b} log p_k = \sum_k f_k \frac{1}{k}\frac{p_{k-1}}{p_k} + \sum_k f_k \frac{\partial}{\partial b} log p_{k-1}$$

Let

$$S_b = \sum_k f_k \frac{\partial}{\partial b} log p_k$$

Then

$$S_b = \sum_k f_k \frac{1}{k}\frac{p_{k-1}}{p_k} + \sum_k f_k \frac{\partial}{\partial b} log p_{k-1} \qquad (5.12)$$

The Maximum Likelihood equations are $S_a = 0$ and $S_b = 0$ i.e.

$$S_a = \sum_k f_k \frac{p_{k-1}}{p_k} + \sum_k f_k \frac{\partial}{\partial a} log p_{k-1} = 0$$

$$S_b = \sum_k f_k \frac{1}{k} \frac{p_{k-1}}{p_k} + \sum_k f_k \frac{\partial}{\partial b} log p_{k-1} = 0$$

which cannot be simplified further. This shows that this method does not work when applied to the Panjer pattern.

5.1.2 Estimation based on Compound Distributions (Poisson Binomial Distribution)

The Poisson Binomial distribution is

$$p_k = e^{-a} \sum_{t=0}^{\infty} \frac{a^t}{t!} \binom{nt}{k} p^k q^{nt-k}$$

where n is a known integer and $q = 1 - p$.
Then p_k obeys the Recursion formula

$$p_{k-1} = e^{-a} \sum_{t=0}^{\infty} \frac{a^t}{t!} \binom{nt}{k-1} p^{k-1} q^{nt-(k-1)}$$

Let

$$S_1(k) = \sum_{t=0}^{\infty} \frac{a^t}{t!} \binom{nt}{k} p^k q^{nt-k}$$

$$\Longrightarrow p_k = e^{-a} S_1(k)$$

Hence

$$S_1(k) = e^a p_k$$

Let

$$S_2(k) = \frac{\partial S_1(k)}{\partial a} = \frac{\partial}{\partial a} \sum_{t=0}^{\infty} \frac{a^t}{t!} \binom{nt}{k} p^k q^{nt-k}$$

$$= \sum_{t=0}^{\infty} t \frac{a^{t-1}}{t!} \binom{nt}{k} p^k q^{nt-k}$$

Substitute $t = \frac{nt-k+k}{n}$ **in** $S_2(k)$ **i.e.**

$$S_2(k) = \sum_{t=0}^{\infty} \frac{(nt-k+k)a^{t-1}}{n(t!)} \binom{nt}{k} p^k q^{nt-k}$$

$$= \sum_{t=0}^{\infty} \frac{(nt-k)a^{t-1}}{n(t!)} \binom{nt}{k} p^k q^{nt-k} + \sum_{t=0}^{\infty} k \frac{a^{t-1}}{n(t!)} \binom{nt}{k} p^k q^{nt-k}$$

$$= \sum_{t=0}^{\infty} \frac{a^{t-1}}{n(t!)} \frac{(k+1)(nt)!}{(k+1)!(nt-(k+1))!} p^k q^{nt-k} + \sum_{t=0}^{\infty} k \frac{a^{t-1}}{n(t!)} \binom{nt}{k} p^k q^{nt-k}$$

$$= \frac{(k+1)}{n} \sum_{t=0}^{\infty} \frac{a^{t-1}}{t!} \binom{nt}{k+1} p^{k+1} q^{nt-k-1} \frac{q}{p} + \frac{k}{n} \sum_{t=0}^{\infty} \frac{a^{t-1}}{t!} \binom{nt}{k} p^k q^{nt-k}$$

$$= \frac{q(k+1)}{np} \frac{S_1(k+1)}{a} + \frac{k}{n} \frac{S_1(k)}{a}$$

Note:

$$(nt-k)\binom{nt}{k} = (nt-k)\frac{(nt)!}{k!(nt-k)!} = \frac{(nt)!}{k!(nt-k-1)!} = (k+1)\frac{(nt)!}{(k+1)k!(nt-k-1)!} = (k+1)($$

Therefore

$$S_2(k) = \frac{q(k+1)}{nap} S_1(k+1) + \frac{k}{na} S_1(k) \qquad (5.13)$$

But

$$S_1(k) = e^a p_k$$

$$\Longrightarrow S_1(k+1) = e^a p_{k+1}$$

Substitute in $S_2(k)$ **then**

$$S_2(k) = \frac{q(k+1)}{np} \frac{S_1(k+1)}{a} + \frac{k}{n} \frac{S_1(k)}{a} \qquad (5.14)$$

Further let

$$
\begin{aligned}
S_3(k) &= \frac{\partial S_1(k)}{\partial p} = \frac{\partial}{\partial p}\sum_{t=0}^{\infty}\frac{a^t}{t!}\binom{nt}{k}p^k q^{nt-k} \\
&= \sum_{t=0}^{\infty}\frac{a^t}{t!}\binom{nt}{k}[kp^{k-1}q^{nt-k} - p^k(nt-k)(1-p)^{nt-k-1}] \\
&= \sum_{t=0}^{\infty}\frac{a^t}{t!}\binom{nt}{k}kp^{k-1}q^{nt-k} - \sum_{t=0}^{\infty}\frac{a^t}{t!}(nt-k)\binom{nt}{k}p^k q^{nt-k-1} \\
&= k\sum_{t=0}^{\infty}\frac{a^t}{t!}\binom{nt}{k}p^{k-1}q^{nt-k} - (k+1)\sum_{t=0}^{\infty}\frac{a^t}{t!}\binom{nt}{k+1}p^k q^{nt-k-1} \\
&= k\sum_{t=0}^{\infty}\frac{a^t}{t!}\binom{nt}{k}p^k q^{nt-k}\frac{1}{p} - (k+1)\sum_{t=0}^{\infty}\frac{a^t}{t!}\binom{nt}{k+1}p^{k+1}q^{nt-k-1}\frac{1}{p} \\
&= \frac{k}{p}S_1(k) - \frac{k+1}{p}S_1(k+1)
\end{aligned}
$$

Therefore

$$
S_3(k) \;=\; \frac{k}{p}S_1(k) - \frac{k+1}{p}S_1(k+1) \tag{5.15}
$$

OR

$$
S_3(k) \;=\; \frac{k}{p}e^a p_k - \frac{k+1}{p}e^a p_{k+1} \tag{5.16}
$$

Now since

$$
\begin{aligned}
p_k &= e^{-a}S_1(k) \\
\Longrightarrow logp_k &= -a + S_1(k)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial a}logp_k &= -1 + \frac{\partial}{\partial a}S_1(k) \\
&= -1 + \frac{S_2(k)}{S_1(k)} \\
&= -1 + \frac{\frac{q(k+1)}{nap}S_1(k+1) + \frac{k}{na}S_1(k)}{S_1(k)} \\
&= -1 + \frac{q(k+1)}{nap}\frac{S_1(k+1)}{S_1(k)} + \frac{k}{na}
\end{aligned}
$$

But

$$\frac{S_1(k+1)}{S_1(k)} = \frac{p_{k+1}}{p_k}$$

Therefore

$$\frac{\partial}{\partial a} log p_k = -1 + \frac{q(k+1)}{nap}\frac{p_{k+1}}{p_k} + \frac{k}{na} \qquad (5.17)$$

$$\begin{aligned}\frac{\partial}{\partial p} log p_k &= -1 + \frac{\partial}{\partial p} S_1(k) \\ &= \frac{S_3(k)}{S_1(k)} \\ &= \frac{\frac{k}{p}S_1(k) - \frac{k+1}{p}S_1(k+1)}{S_1(k)} \\ &= \frac{k}{p} - \frac{k+1}{p}\frac{S_1(k+1)}{S_1(k)}\end{aligned}$$

$$\frac{\partial}{\partial p} log p_k = \frac{k}{p} - \frac{k+1}{p}\frac{p_{k+1}}{p_k} \qquad (5.18)$$

Multiply equation(5.17) and (5.18) by a_k and sum over k,

$$\begin{aligned}\sum_k a_k \frac{\partial}{\partial a} log p_k &= -\sum_k a_k + \sum_k a_k \frac{q(k+1)}{nap}\frac{p_{k+1}}{p_k} + \sum_k a_k \frac{k}{na} \\ &= -n + \frac{q}{nap}\sum_k a_k(k+1)\frac{p_{k+1}}{p_k} + \frac{1}{a}\frac{\sum_k k a_k}{n} \\ &= -n + \frac{q}{nap}\sum_k a_k(k+1)\frac{p_{k+1}}{p_k} + \frac{\bar{k}}{a}\end{aligned}$$

$$\begin{aligned}\sum_k a_k \frac{\partial}{\partial p} log p_k &= \sum_k a_k \frac{k}{p} - \sum_k a_k \frac{k+1}{p}\frac{p_{k+1}}{p_k} \\ &= \frac{1}{p}\sum_k k a_k - \frac{1}{p}\sum_k (k+1) a_k \frac{p_{k+1}}{p_k} \\ &= \frac{1}{p} n\bar{k} - \frac{1}{p}\sum_k (k+1) a_k \frac{p_{k+1}}{p_k}\end{aligned}$$

since $\bar{k} = \frac{\sum_k k a_k}{\sum_k a_k} = \frac{\sum_k k a_k}{n}$

Let

$$S_a = \sum_k a_k \frac{\partial}{\partial a} log p_k$$

Then

$$S_a = -n + \frac{q}{nap}\sum_k (k+1)a_k \frac{p_{k+1}}{p_k} + \frac{\bar{k}}{a} \tag{5.19}$$

Let

$$S_p = \sum_k a_k \frac{\partial}{\partial p} log p_k$$

Then

$$S_p = \frac{n\bar{k}}{p} - \frac{1}{p}\sum_k (k+1)a_k \frac{p_{k+1}}{p_k} \tag{5.20}$$

The Maximum Likelihood equations are $S_a = 0$ and $S_p = 0$ i.e.

$$S_a = -n + \frac{q}{nap}\sum_k (k+1)a_k \frac{p_{k+1}}{p_k} + \frac{\bar{k}}{a} = 0 \tag{5.21}$$

$$S_p = \frac{n\bar{k}}{p} - \frac{1}{p}\sum_k (k+1)a_k \frac{p_{k+1}}{p_k} = 0 \tag{5.22}$$

where a_k is the observed frequency of k and $n = \sum_k a_k$,the total number of observations.

From equation(5.22),

$$pS_p = n\bar{k} - \sum_k (k+1)a_k \frac{p_{k+1}}{p_k} = 0$$

$$\Longrightarrow \sum_k (k+1)a_k \frac{p_{k+1}}{p_k} = n\bar{k} - pS_p$$

Substituting this value in equation(5.21) one gets

$$S_a = -n + \frac{q}{nap}[n\bar{k} - pS_p] + \frac{\bar{k}}{a} = 0$$

i.e.

$$-n+\frac{q}{nap}[n\bar{k}-pS_p]+\frac{\bar{k}}{a}=0$$

but $S_p = 0$ then

$$-n+\frac{qn\bar{k}}{nap}+\frac{\bar{k}}{a}=0$$

$$-n+\frac{(1-p)n\bar{k}}{nap}+\frac{\bar{k}}{a}=0$$

$$\Longrightarrow -n+\frac{\bar{k}}{ap}=0$$

$$\bar{k} \;=\; n\hat{a}\hat{p} \tag{5.23}$$

where $\hat{a}$ and $\hat{p}$ are the maximum likelihood estimates of a and p respectively.

Again from equation(5.22),

$$n\bar{k}-\sum_k (k+1)a_k\frac{p_{k+1}}{p_k}=0$$

$$\Longleftrightarrow n\bar{k}=\sum_k (k+1)a_k\frac{p_{k+1}}{p_k}$$

$$\Longrightarrow \bar{k}=\frac{1}{n}\sum_k (k+1)a_k\frac{p_{k+1}}{p_k}$$

Substituting this value of $\bar{k}$ in equation(5.21) one gets

$$\begin{aligned} S_a &= -n+\frac{q}{nap}\sum_k (k+1)a_k\frac{p_{k+1}}{p_k}+\frac{1}{na}\sum_k (k+1)a_k\frac{p_{k+1}}{p_k} \\ &= -n+\frac{1-p}{nap}\sum_k (k+1)a_k\frac{p_{k+1}}{p_k}+\frac{1}{na}\sum_k (k+1)a_k\frac{p_{k+1}}{p_k} \\ &= -n+\frac{1}{nap}\sum_k (k+1)a_k\frac{p_{k+1}}{p_k} \end{aligned}$$

$$\Longrightarrow S_a=\sum_k a_k\frac{\partial}{\partial a}logp_k=-n+\frac{1}{nap}\sum_k (k+1)a_k\frac{p_{k+1}}{p_k}$$

i.e.

$$\sum_k a_k \frac{\partial}{\partial a} log p_k = \frac{1}{nap} \sum_k (k+1) a_k \frac{p_{k+1}}{p_k} - n$$

Therefore

$$L(\bar{p}) \;\; = \;\; \sum_k a_k F(k) - n \tag{5.24}$$

where

$$L(\hat{p}) = \sum_k a_k \frac{\partial}{\partial a} log p_k$$

and

$$F(k) = \frac{(k+1) p_{k+1}}{n\bar{a}\bar{p} p_k}$$

Chapter 6

CONCLUSION AND RECOMMENDATIONS

6.1 Objectives

The objectives of this project are:-

a) To construct discrete probability distributions.
Specifically to construct
-Standard discrete distributions
-Compound discrete distributions

b) To express pmfs in terms of recursive relations specifically expressing p_k in terms of p_{k-1} where $p_k = Prob(X = k)$

c) To study the recursive relations by
(i)Identifying patterns and corresponding distributions.
(ii)Deriving estimation procedures for recursive relations

6.2 What has been achieved

From Chapter 2,we have the following:-

- From the exponential expansion we obtained;
-Poisson Distribution
-Zero-truncated Poisson Distribution

- From the Binomial expansion we obtained;
  -Binomial Distribution
  -Bernoulli Distribution
  -Negative Binomial Distribution-2cases
  -Geometric Distribution-2cases
  -Logarithmic Distribution

- Using Change of Variable(Jacobian Transformation) technique we obtained;
  -Exponential Distribution
  -Gamma Distribution with two parameters
  -Beta density function.

- Using the Expectation and Convolution Approaches to the sums of identically independent distributed (iid) random variables we proved that:
  -the distribution of the sum of Bernoulli iid random variables is a Binomial distribution.
  -the distribution of the sum of Binomial iid random variables is a Binomial distribution.
  -the distribution of the sum of Geometric iid random variables is a Negative Binomial distribution.
  -the distribution of the sum of Negative Binomial iid random variables is another Negative Binomial distribution.
  -the distribution of the sum of Poisson iid random variables is another Poisson distribution.

From Chapter 3: we have derived recursive ratios and recursive relations for various probability distributions that have been identified in Chapter 2.

Using the recursive relations we have derived means and variances,based on pgf technique(where possible) and Feller's method.

From Chapter 4: we have a number of patterns of recursive relations;the basic one being the Panjer pattern of 1981 which is in the form $p_n = \frac{an+b}{n} p_{n-1}; n = 1, 2, \ldots$.With this pattern we have been able to get the corresponding probability distributions,namely;Poisson,Binomial and Negative Binomial distributions.
The probability distributions corresponding to the Willmot (1988)

pattern which is in the form

$$p_n = (a + \frac{b}{n})p_{n-1} \quad ; n = 2, 3, \ldots$$

are:Zero-Truncated Poisson,Zero-Truncated Geometric and Logarithmic distributions.

For estimation,we have applied Sprott's procedure for deriving ML equations [which have to be in a recursive form].

McGilchrist(1969) derived an estimation procedure from the recurrence relationship between probabilities in a discrete distribution.It is limited to cases in which only one parameter occurs in the recurrence relationship.

If the recurrence relation is of the form

$$\beta_0(x,\alpha)p_x + \beta_1(x,\alpha)p_{x-1} = 0$$

where β_0 and β_1 are fuctions of x and the unknown parameter α,then the estimator is obtained from the set of equations

$$\sum_{x=1}^{b} \omega_x[\beta_0(x,\alpha)f_x + \beta_1(x,\alpha)f_{x-1}] = 0$$

where a is an estimator of α,f_x is the frequency of occurrence of x in n independent observations and ω_x are arbitrary weights.but McGilchrist(1969) proposed a suitable criterion for the choice of optimum weights. Examples have been given for Poisson Distribution,Truncated Negative Binomial Distribution and Logarithmic Series Distribution.

6.3 Application

This work is applicable in three areas of research:

- In insurance we can apply the work for claims of premiums.
- In group testing problems we can apply the work in estimation problems.
- In demography we can study Birth Interval Analysis using this work.

6.4 Recommendations

a) Other methods for constructing probability distributions need to be looked at and their parameters estimated.
For example we have estimation of mixtures of discrete distributions.

Let $\{f_i\}$ and $\{g_i\}$ be two probability distributions,$\alpha > 0, \beta > 0, \alpha + \beta = 1$.Then $\{\alpha f_i + \beta g_i\}$ is again a probability distribution.

$\{f_i\}$ and $\{g_i\}$ could be mixed Poisson distributions with parameters $\lambda_1 > \lambda_2$;mixed truncated Poisson distributions with parameters $\lambda_1 > \lambda_2$; and mixed Poisson with parameter λ and Binomial parameters n and p.

The problem is to estimate $\alpha, \beta = 1 - \alpha$ and the parameters of each distribution in the mixture.

b) Further study is required for distributions whose pgfs can be achieved but whose explicit probabilities and hence recursive relations are not simple.

We have also cases of differential equations in pgfs which are not easily solvable.Numerical methods or other methods need to be studied.

c) (i)For Panjer's pattern,we identified the corresponding distributions,namely,Poisson,Binomial and Negative Binomial Distribution using both pgf and iteration technique.
It will be interesting to identify probability distributions for other patterns.
(ii)Sprott(1958) obtained ML equations for Poisson-Binomial distribution,written in the form

$$\sum_k (k+1)a_k \frac{p_{k+1}}{p_k} = N\bar{k}$$

and

$$N\hat{a}\hat{p} = \bar{k}$$

where a_k is the observed frequency of k,$\hat{a}$,$\hat{p}$ are ML estimates of parameters a and p;$\bar{k}$ is the sample mean and N the sample size. The p_k are calculated recursively. It will be interesting to find out ML equations for other distributions.

Bibliography

Feller,W (1968):"An Introduction to Probability Theory and Its Applications" Vol 1,John Wiley,3rd Edition.

Karlin,S and H.M.Taylor (1975):"A First Course in Stochastic Processes Academic Press",2nd Edition.

DePril,N. (1985):"Recursions for Convolutions of Arithmetic Distributions" ASTIN Bulletin Vol 15,No. 2,135-139.

Hesselager,Ole (1994):"A Recursive Procedure for Calculation of some Compound Distributions" ASTIN Bulletin Vol 24,No. 1,19-32.

Panjer,H.H. and S.Wang (1995):"Computational Aspects of Sundt's Generalized class" ASTIN Bulletin Vol 25,No. 1,5-17.

Hess,K.T.;A.Liewald and K.D. Schmidt (2002):"An Extension of Panjer's Recursion" ASTIN Bulletin Vol 32,No. 2,283-297.

Panjer H.H. (1981):"Recursive Evaluation of A Family of Compound Distributions" ASTIN Bulletin Vol 12,22-26.

Wang S. and M.Sobrero (1994):"Further Results on Hesselager's Recursive Procedure for Calculation of some Compound Distributions" ASTIN Bulletin Vol 24,No. 4,161-166

Sundt,B. and W.S.Jewell (1981):"Further Results on Recursive Evaluation of Compound Distributions" ASTIN Bulletin,Vol 12,27-39.

DePril,N. (1986):"Moments of a class of compound distributions" Scandinavian Acturial Journal,117-120.

Guldberg,A. (1931):"On discontinuous frequency-functions and statistical series" Skandanrisk Aktuarietdskrift,14,167-187.

Ord,J.K. (1967):"On a system of discrete distributions" Biometrika 54,649-65.

Panjer,H.H. and Willmot,G.E (1982):"Recursions for Com-

pound Distributions" ASTIN Bulletin Vol 13,1-11.
Schroter,K.J. (1990):"On a Family of counting distributions and recursions for related compound distributions" Scandinavian Acturial Journal(1990),161-174.
Sundt,B. (1992):"On some Extensions of Panjer's class of counting distributions" ASTIN Bulletin Vol 22,61-80.
Willmot,G.E. and Panjer,H.H. (1987):"Difference Equation approaches in evaluation of compound distributions" Insurance Mathematics and Economics 6,43-56.
Elvers,E. (1991):"A note on the generalized Poisson Distribution" Astin Bulletin Vol 21,167.
Goovaerts,M.J. and Kaas R. (1991):"Evaluating compound generalized Poisson distributions recursively" ASTIN Bulletin Vol 21,193-198.
Johnson,N.L.;Kotz,S. and Kemp,A.W. (1992):"Univariate Discrete Distributions" 2nd Edition,John Wiley.
Ruohonen,M. (1988):"A model for the claim process" ASTIN Bulletin Vol 18,57-68.
Willmot,G.E. (1987):"The Poisson -Inverse-Gaussian as an alternative to the Negative Binomial" Scandinavian Acturial Journal,113-127.
Willmot,G.E. (1988):"Sundt and Jewell's family of discrete distributions" ASTIN Bulletin Vol 18,17-29.
Willmot,G.E. (1989):"Limiting tail behaviour of some discrete compound distributions" Insurance Mathematics and Economics 8,175-185.
Willmot,G.E. (1993):"On the recursive evaluation of mixed Poisson probabilities and related quantities" Scandinavian Acturial Journal,Vol 2,114-133.
Willmot,G.E. and Sundt,B.(1989):"On Evaluation of the Delaporte distribution and related distributions" Scandinavian Acturial Journal,101-113.
Panjer H.H. and S. Wang(1993):"On the stability of recursive formulas" ASTIN Bulletin Vol 23,No 2,227-258.
Willmot,G.E. (1986):"Mixed Compound Poisson Distribution" ASTIN Bulletin Vol 16,No 5,59-76.