

# DESCRIPTION

This document provides solution's to 3.1 to 3.15 from chapter 3 of statistical inference book by Casella and Berger, focusing on discrete probability distribution

U.S.G. 7.1

3.1

$$X \sim \text{Unif}(N_0, N_1)$$

$$\frac{1}{N_1 - N_0 + 1}$$

where  $X = N_0, N_1, \dots, N_1$

$N_0 \leq N_1$  & integer's

$E[X], \text{Var}(X) ?$

$$P(X=x) = \begin{cases} \frac{1}{N_1 - N_0 + 1} & \text{if } x \in \{N_0, N_1, \dots, N_1\} \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \sum_{x=N_0}^{N_1} x \cdot P(X=x)$$

$$= \frac{1}{N_1 - N_0 + 1} \cdot \sum_{x=N_0}^{N_1} x$$

$$= \frac{1}{N_1 - N_0 + 1} \cdot \left( \sum_{x=0}^{N_1} x - \sum_{x=0}^{N_0-1} x \right)$$

$$= \frac{1}{N_1 - N_0 + 1} \cdot \left( \frac{N_1(N_1+1)}{2} - \frac{(N_0-1)N_0}{2} \right)$$

$$= \frac{1}{N_1 - N_0 + 1} \left( \frac{N_1^2 + N_1 - N_0^2 + N_0}{2} \right)$$

$$= \frac{1}{\cancel{N_1 - N_0 + 1}} \frac{(N_1 - \cancel{N_0 + 1})(N_1 + N_0)}{2}$$

$E[X] = \frac{N_1 + N_0}{2}$

$$E[X^2] = \sum_{x=N_0}^{N_1} x^2 P(X=x)$$

$$= \frac{1}{N_1 - N_0 + 1} \left[ \sum_{x=0}^{N_1} x^2 - \sum_{x=0}^{N_0-1} x^2 \right]$$

$$= \frac{1}{N_1 - N_0 + 1} \left[ \frac{N_1(N_1+1)(2N_1+1)}{6} - \frac{(N_0-1)N_0(N_0+1)}{6} \right]$$

$$\text{var}(X) = E[X^2] - E[X]^2$$

= Substituting these we get

$\text{var}(X) = \frac{(N_1 - N_0)(N_1 - N_0 + 2)}{12}$

3.2

A manufacturer receives a lot of 100 parts from a vendor. The lot will be unacceptable if more than 5 of the parts are defective. The manufacturer is going to select randomly  $k$  parts from the lot for inspection and the lot will be accepted if no defective parts are found in the sample.

(a)

How large does  $k$  have to be to ensure that the probability that the manufacturer accepts an unacceptable lot is less than 0.10?

Solution:

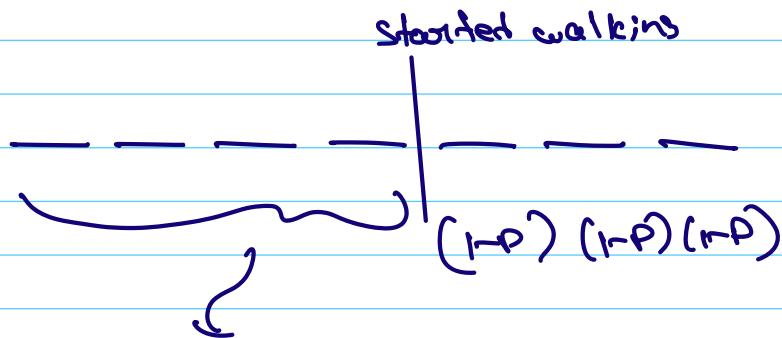
$n = \text{number of defective parts}$

$\Rightarrow$  for a lot to be unacceptable  $n > 6$

3.3

Prob of a car passing during any given second is a const  $P$  and there is no interaction b/w the passing of cars at different seconds.

Solution:



only way that he will wait for 1st 4 seconds is (at least one car pass in 1st 3 seconds i.e.  $(1-P^3)$ ) and a car pass in 4th second.

$$\Rightarrow (1-P^3) P (1-P)$$

3.4

A man with  $n$  keys wants to open his door and tries the key at random. Exactly one key will open the door. Find the mean Number of trials if

a) unsuccessful key's are not eliminated from further selection's.

Solution:

IP (selecting right key out of  $n$ ) =  $\frac{1}{n} = p$

occurrence of 1st success without eliminating the unsuccessful key's

$X = \# \text{ of trials}$

( $x-1$ ) unsuccessful,  $x^{\text{th}}$  trial success

$X \sim \text{geometric}(p)$

$$\begin{aligned} \text{IP}(X=x) &= (1-p)^{x-1} \cdot p \\ &= \left(\frac{99}{100}\right)^{x-1} \cdot p \end{aligned}$$

$$\mathbb{E}[X] = \frac{1}{p} = 100$$

b)

unsuccessful keys are eliminated

$X =$  number trials

= ( $x-1$ ) unsuccessful ,  $x^{th}$  success

$$P(X=1) = \frac{1}{100} = \frac{1}{100}$$

$$P(X=2) = \frac{99}{100} \cdot \frac{1}{99} = \frac{1}{100}$$

$$P(X=3) = \frac{99}{100} \cdot \frac{98}{99} \cdot \frac{1}{98} = \frac{1}{100}$$

$$P(X=4) = \frac{99}{100} \cdot \frac{98}{99} \cdot \frac{97}{98} \cdot \frac{1}{97} = \frac{1}{100}$$

$$\Rightarrow P(X=x) = \frac{1}{100} \quad x \in \{1, 2, \dots, 100\}$$

$$E[X] = \frac{N+1}{2} = \frac{101}{2} = 50.5$$

3.5

A standard drug is known to be effective in 80% of the cases in which it is used. A new drug is tested on 100 patients and found to be effective in 85 cases. Is the new drug superior?

Solution:

$$P(\text{success when tested on a patient}) = 0.8$$

$$X \sim \text{Binomial}(100, 0.8)$$

Where  $x$  is number of success when tested on 100 patients

then  $P(X > 85) =$

$$\sum_{x=85}^{100} \binom{100}{x} (0.8)^x (0.2)^{100-x}$$

$$P(X > 85) = 0.1285$$

this Probability is for old drug.

\* the Probability that the old drug works on more than 85 = 0.1285

$\Rightarrow$  which means there is a significant (0.1285 is not small) time's old works on 85 patients

$\Rightarrow$  if the new drug work on 85 patients out of 100, we cannot conclude that it is good

3.6

A Large number of insects are expected to be attracted to a certain variety of plants. A commercial insecticide is advertised as being 99% effective. Suppose 2000 insects infest a rose garden where the insecticide

has been applied, and  $X =$  number of surviving insects.

- (a) What Prob dist might provide a reasonable model for this experiment?

$$X \sim \text{Binomial}(2000, \frac{1}{100})$$

- (b) Prob that fewer than 100 insects survive

$$P(X \leq 100) = \sum_{x=0}^{100} \binom{2000}{x} \left(\frac{1}{100}\right)^x \left(\frac{99}{100}\right)^{2000-x}$$

- (c) evaluate an approximation to the probability in part (b)

Solution:

out of 2000, Expected number of surviving =  $E[X] = np = 20$

We can approximate this with Poisson  
with  $\lambda = 20$

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

$$\Rightarrow P(X \leq 100) = \sum_{x=0}^{100} \frac{e^{-20} \cdot 20^x}{x!}$$
$$= 0.99$$

3.7

Let the number of chocolate chips in a certain type of cookie have a Poisson distribution. We want the probability that a randomly chosen cookie has at least two chocolate chips to be greater than 0.99. Find the smallest value of the mean of the distribution that ensures this probability.

Solution:

$\text{P}(\text{randomly chosen cookie has at least two chocolate chips}) \geq 0.99$

$X \sim \text{Number of chocolate chips on cookies}$

$$\text{P}(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

$$\text{P}(X \geq 2) \geq 0.99$$

$\Rightarrow$  Probability that atleast 2 chocolate chips on a cookie

$$\Rightarrow 1 - \text{P}(X=0) - \text{P}(X=1) \geq 0.99$$

$$\Rightarrow 1 - e^{-\lambda} - e^{-\lambda} \cdot \lambda \geq 0.99$$

$$\Rightarrow 0.01 \geq e^{-\lambda} \cdot (1 + \lambda)$$

Using tail end error

$$\lambda = 6.62825$$

3.8

a)

Probability of turning away a customer (because of a full house)

is less than 1%

Solution:

Probability of a person coming to

Choosing a theater =  $\frac{1}{2}$

$X \sim \text{Binomial}(1000, \frac{1}{2})$

$$P(X > N) = \sum_{x=N+1}^{1000} \binom{1000}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{1000-x} < 0.01$$

$$\Rightarrow \sum_{x=N+1}^{1000} \binom{1000}{x} \left(\frac{1}{2}\right)^{1000} < 0.01$$

Solve this inequality. ( $N = 537$  around)

b)

$X \sim \text{Binomial}(1000, \frac{1}{2})$

$$\begin{aligned} E[X] &= np = 500 & \text{Var}(X) &= np(1-p) \\ & & &= 250 \end{aligned}$$

$Y \sim N(500, 250)$

$$Y \sim N(500, \sqrt{250})$$

$$P(X > N) < 0.01$$

$$\Rightarrow P(\sqrt{250} Z + 500 > N) < 0.01$$

$$\Rightarrow P(Z > \frac{N - 500}{\sqrt{250}}) < 0.01$$

from Normal table

$$P(Z > 2.33) \approx 0.099 < 0.01$$

$$\Rightarrow \frac{N - 500}{\sqrt{250}} = 2.33$$

$$\Rightarrow$$

$$N = 537$$

3.9

a)  $X \sim \text{Binomial}(60, \frac{1}{90})$

$$P(X \geq s) = \sum_{x=s}^{60} \binom{60}{x} \left(\frac{1}{90}\right)^x \left(\frac{89}{90}\right)^{60-x}$$
$$= 0.006$$

(b) 62 Counties, and 5 schools in  
each County

$$\Rightarrow 62 \times 5 = 310$$

$X \sim$  Number of School's with 5 or more  
twin's

$$X \sim \text{Binomial}(310, 0.0006)$$

then  $P(X \geq 1) = 1 - P(X=0)$

$$= 1 - \binom{310}{0} \left(\frac{89}{90}\right)^{310}$$

3.12

$X \sim \text{binomial}(n, p)$

$Y \sim \text{negative binomial}(r, p)$

Show that  $F_X(r-1) = 1 - F_Y(n-r)$

Solution:

$$F_X(r-1) = \sum_{x=0}^{r-1} \binom{n}{x} p^x (1-p)^{n-x}$$

$$F_Y(n-r) = \sum_{y=0}^{n-r} \binom{r+y-1}{y} p^r (1-p)^y$$

$$1 - F_X(r-1) = \sum_{x=r}^n \binom{n}{x} p^x (1-p)^{n-x}$$

3.13

A truncated discrete distribution is one in which a particular class cannot be observed and is eliminated from the sample space. In particular, if  $X$  has range  $0, 1, 2, \dots$  and the 0 class cannot be observed (as is usually the case), the 0-truncated r.v.  $X_T$  has

Pmf

$$P(X_T = x) = \frac{P(X=x)}{P(X>0)}, x \in 1, 2, \dots$$

Find the Pmf, mean, and variance of the 0-truncated r.v.

$$\begin{aligned} P(X_T = x) &= P(X=x | X>0) \\ &= \frac{P(X=x \cap X>0)}{P(X>0)} \end{aligned}$$

The event  $X=x$  and  $X>0$  will have the intersection with  $X=x$

i.e

$$P(X=x \cap X > 0) = P(X=x)$$

therefore

$$P(X_T=x) = P(X=x | X > 0)$$

$$= \frac{P(X=x \cap X > 0)}{P(X > 0)}$$

$$= \frac{P(X=x)}{P(X > 0)}$$

②  $X \sim \text{Poisson}(\lambda)$

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \quad x=0, 1, 2, \dots$$

$$P(X_T=x) = \frac{P(X=x)}{P(X > 0)} \quad x=1, 2, \dots$$

$$= \frac{\frac{e^{-\lambda} \cdot \lambda^x}{x!}}{1 - e^{-\lambda}} \quad (P(X > 0) = 1 - P(X=0))$$

PMF

$$P(X_T=x) = \frac{e^{-\lambda}}{1-e^{-\lambda}} \cdot \frac{\lambda^x}{x!}$$
$$x = 1, 2, \dots$$

$$E[X_T] = \sum_{x=1}^{\infty} x \cdot P(X_T=x)$$

$$= \sum_{x=1}^{\infty} x \cdot \frac{P(X=x)}{P(X>0)}$$

$$= \frac{1}{P(X>0)} \sum_{x=1}^{\infty} x \cdot P(X=x)$$

$$= \frac{1}{P(X>0)} \cdot E[X] = \frac{\lambda}{1-e^{-\lambda}}$$

$$E[X] = \frac{\lambda}{1-e^{-\lambda}}$$

$$E[X_T(X_{T-1})] = \sum_{x=1}^{\infty} \frac{x(x-1) \cdot P(X=x)}{P(X>0)}$$

$$= \frac{1}{\text{IP}(X>0)} \sum_{x=1}^{\infty} x(x-1) \text{IP}(X=x)$$

$$= \frac{1}{\text{IP}(X>0)} \sum_{x=2}^{\infty} \frac{x(x-1) e^{-\lambda} \cdot \lambda^x}{x!}$$

$$= \frac{\lambda^2}{1-e^{-\lambda}} \cdot \sum_{x=2}^{\infty} \underbrace{\frac{e^{-\lambda} \cdot \lambda^{x-2}}{(x-2)!}}_{\text{PMF} = 1}$$

$$\mathbb{E}[X_T(X_{T-1})] = \frac{\lambda^2}{1-e^{-\lambda}}$$

$$\Rightarrow \text{Var}(X_T) = \mathbb{E}[X_T(X_{T-1})]$$

$$+ E[X_T] - E[X_T]^2$$

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

$$= \frac{1}{(1-e^{-\lambda})^2} \left[ \lambda^2(1-e^{-\lambda}) + \lambda(1-e^{-\lambda}) - \lambda^2 \right]$$

$$= \frac{1}{(1-e^{-\lambda})^2} \left[ -\lambda^2 e^{-\lambda} + \lambda - \lambda e^{-\lambda} \right]$$

$$\text{Var}(X_T) = \frac{\lambda}{(1-e^{-\lambda})^2} - \frac{e^{-\lambda} (\lambda^2 + \lambda)}{(1-e^{-\lambda})^2}$$

(b)

$X \sim \text{negative binomial } (\alpha, p)$

$$\text{IP}(X=x) = \binom{\alpha+x-1}{x} p^\alpha (1-p)^x$$

$$x = 0, 1, 2, \dots$$

Solution:

$$\text{IP}(X=0) = \binom{\alpha-1}{0} p^\alpha$$

$$\text{IP}(X>0) = 1 - \text{IP}(X=0)$$

$$= 1 - p^\alpha$$

$$\Rightarrow \text{IP}(X_T=x) = \frac{\text{IP}(X=x)}{\text{IP}(X>0)}$$

PMF

$$P(X_T=x) = \frac{1}{1-p\alpha} \binom{\alpha+x-1}{x} p^x (1-p)^{\alpha}$$

$x=1, 2, \dots$

$$E[X_T] = \frac{1}{P(X>0)} \sum_{x=1}^{\infty} x \cdot P(X=x)$$

$$= \frac{E[x]}{P(X>0)} = \frac{1}{1-p\alpha} \frac{\alpha(1-p)}{p}$$

$\Rightarrow$

$$E[X_T] = \alpha \cdot \frac{1-p}{P(1-p\alpha)}$$

$$E[X_T(1-X_T)] = \sum_{x=2}^{\infty} \frac{P(X=x) \cdot x(1-x)}{P(X>0)}$$

$$= \frac{1}{P(X>0)} \sum_{x=2}^{\infty} x(1-x) P(X=x)$$

$$= \frac{1}{P(X>0)} \cdot E[X(X-1)]$$

$$E[X_T(X_{T-1})] = \frac{1}{P(X>0)} \sum_{x=2}^{\infty} x(x) \binom{x+1-1}{x} p^x (1-p)^{x-1}$$

$$= \frac{1}{P(X>0)} \sum_{x=2}^{\infty} x(x) \cdot \frac{(x+1-1)!}{x! (x-1)!} p^x (1-p)^{x-1}$$

$$= \frac{1}{P(X>0)} \sum_{x=2}^{\infty} \frac{(x+1-1)!}{(x-2)! (x+1)!} p^{x+2} \cdot (1-p)^{x-2} \cdot \frac{x(x+1)(1-p)^2}{p^2}$$

$$= \frac{1}{P(X>0)} \frac{x(x+1)(1-p)^2}{p^2} \cdot \sum_{x=2}^{\infty} \binom{x+1-1}{x-2} p^{x+2} \cdot (1-p)^{x-2}$$

*negative binomial  $(x+2, p)$*

$$= 1$$

$$= \frac{1}{(1-p^x)} \cdot \left(\frac{1-p}{p}\right)^2 \cdot x(x+1)$$

$$Var(X_T) = E[X_T(X_{T-1})] + E[X_T]$$

$$-E[X_T]^2$$

$$= \frac{1}{(1-p^x)} \left(\frac{1-p}{p}\right)^2 \cdot x(x+1) + \frac{1}{(1-p^x)} \cdot \left(\frac{1-p}{p}\right) x$$

$$- \frac{1}{(1-p\alpha)^2} \left( \frac{1-p}{p} \right)^2 \alpha^2$$

$$\text{Var}(X) = \frac{1}{(1-p\alpha)} \cdot \left( \frac{1-p}{p} \right) \cdot \alpha \left[ \frac{(1-p)(\alpha+1)}{p} - 1 \right]$$

$$- \frac{1}{(1-p\alpha)^2} \left( \frac{1-p}{p} \right)^2 \alpha^2$$

3.14

Starting from the or-truncated negative binomial (refer to Exercise 3.13), if we let  $\alpha \rightarrow 0$ , we get an interesting distribution, the logarithmic series distribution. A r.v.  $X$  has a logarithmic series distribution with parameters  $p$  if

$$P(X=x) = \frac{-(1-p)^x}{x \log p}, x=1, 2, \dots \quad 0 < p < 1$$

(a) Verify that this defines a legitimate probability function.

$$P(X=x) = \frac{-(1-p)^x}{x \log p}, \quad x=1, 2, \dots$$

$0 < p < 1$

$$\sum_{x=1}^{\infty} P(X=x) = \sum_{x=1}^{\infty} \frac{-(1-p)^x}{x \log p}$$

$$= \frac{-1}{\log p} \sum_{x=1}^{\infty} \frac{(1-p)^x}{x}$$

Taylor series expansion of  $\log x$

$$\ln(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3$$

$$- \frac{1}{4}(x-1)^4 + \dots$$

here  $\ln(p)$  expansion where  $p < 1$

$$\begin{aligned} \Rightarrow \ln(p) &= (p-1) - \frac{1}{2}(p-1)^2 + \frac{1}{3}(p-1)^3 \\ &\quad - \frac{1}{4}(p-1)^4 + \dots \end{aligned}$$

$$= -(1-p) - \frac{1}{2}(1-p)^2 - \frac{1}{3}(1-p)^3$$

$$- \frac{1}{4}(1-p)^4 - \frac{1}{5}(1-p)^5 - \dots$$

$$= - \sum_{x=1}^{\infty} \frac{(1-p)^x}{x}$$

$$\Rightarrow \sum_{x=1}^{\infty} \text{IP}(X=x) = \frac{-1}{\log p} \sum_{x=1}^{\infty} \frac{(1-p)^x}{x}$$

$$= \frac{-1}{\log p} \cdot -\log p$$

$$= 1$$

$$\Rightarrow \sum_{x=1}^{\infty} \text{IP}(X=x) = 1$$

and

$$\text{IP}(X=x) \geq 0 \quad \forall x=1,2,\dots \infty$$

Therefore this is legitimate PMF

6

Find the mean and variance of  $X$

$$\begin{aligned} E[X] &= \sum_{x=1}^{\infty} \frac{-(1-p)^x}{x \log p} \cdot x \\ &= -\frac{1}{\log p} \sum_{x=1}^{\infty} (1-p)^x \\ &= -\frac{1}{\log p} \left( \sum_{x=0}^{\infty} (1-p)^x - 1 \right) \\ &= -\frac{1}{\log p} \left( \frac{1}{1-(1-p)} - 1 \right) = -\frac{1}{\log p} \left( \frac{1-p}{p} \right) \end{aligned}$$

$$E[X] = -\frac{1}{\log p} \left( \frac{1-p}{p} \right)$$

$$\begin{aligned} E[X^2] &= -\frac{1}{\log p} \sum_{x=1}^{\infty} \frac{x^2 \cdot (1-p)^x}{x} \\ &= -\frac{1}{\log p} \sum_{x=1}^{\infty} x \cdot (1-p)^x \end{aligned}$$

identity:

$$\sum_{k=0}^{\infty} a z^k = \frac{a}{1-z} \quad \text{if } |z| < 1$$

Now differentiate both side's with  $\sigma$

$$\frac{d}{d\sigma} \sum_{k=0}^{\infty} a \sigma^k = \frac{d}{d\sigma} \frac{a}{1-\sigma}$$

$$\Rightarrow \sum_{k=0}^{\infty} k \cdot a \sigma^{k-1} = \frac{a}{(1-\sigma)^2}$$

for  $k=0$  the summation = 0

$$\Rightarrow \sum_{k=1}^{\infty} a \cdot k \cdot \sigma^{k-1} = \frac{a}{(1-\sigma)^2}$$

$$E[X^2] = \frac{1}{\log p} \sum_{x=1}^{\infty} x (1-p)^x = \frac{(1-p)}{\log p} \left( \sum_{x=1}^{\infty} x (1-p)^{x-1} \right)$$

$$= -\frac{(1-p)}{\log p} \cdot \frac{1}{p^2}$$

$$\text{Var}(X) = E[X^2] - E[X]^2$$

$$= -\frac{(1-p)}{\log p} \cdot \frac{1}{p^2} - \frac{1}{(\log p)^2} \frac{(1-p)^2}{p^2}$$

$$= \frac{-(1-p)}{\log p \cdot p} = \left( \frac{1}{p} + \frac{1-p}{\log p} \right)$$

$$\text{Var}(X) = -\frac{(1-p)}{p} \cdot \frac{1}{\log p} \left( \frac{1}{p} + \frac{1-p}{\log p} \right)$$

3.15

In section 3.2 it was claimed that the Poisson( $\lambda$ ) distribution is the limit of the negative Binomial ( $\text{NB}_1, p$ ) distribution as  $n \rightarrow \infty$ ,  $p \rightarrow 1$ , and  $n(1-p) \rightarrow \lambda$ .

Show that under these conditions the pmf of the negative binomial converges to that of the Poisson.

Solution:

$$X \sim \text{Poisson}(\lambda), M_X(t) = e^{\lambda(e^t - 1)}$$

$X \sim \text{negative Binomial } (\text{NB}_1, p)$

$$\text{IP}(X=x) = \binom{\text{NB}+x-1}{x} p^{\text{NB}} (1-p)^{x-1}$$

$x = 0, 1, 2, \dots$

$$M_x(t) = E[e^{tx}]$$

$$= \sum_{x=0}^{\infty} e^{tx} \binom{s_1+x-1}{x} p^{s_1} (1-p)^x$$

$$= \sum_{x=0}^{\infty} \binom{s_1+x-1}{x} p^{s_1} (e^{t(1-p)})^x$$

$$= \sum_{x=0}^{\infty} \binom{s_1+x-1}{x} p^{s_1} (1 - (e^{t(1-p)}))^{s_1} [e^{t(1-p)}]^x$$

$$= \frac{p^{s_1}}{[1 - e^{t(1-p)}]^{s_1}} \sum_{x=0}^{\infty} \binom{s_1+x-1}{x} (1 - e^{t(1-p)})^{s_1} (e^{t(1-p)})^x$$

negative binomial  $(s_1, e^{t(1-p)}) = 1$

$$M_x(t) = \frac{p^{s_1}}{[1 - e^{t(1-p)}]^{s_1}}$$

$$\lim_{\begin{matrix} \alpha \rightarrow \infty \\ p \rightarrow 1 \\ \alpha(1-p) \rightarrow \lambda \end{matrix}} M_x(t) = \lim_{\begin{matrix} \alpha \rightarrow \infty \\ p \rightarrow 1 \\ \alpha(1-p) \rightarrow \lambda \end{matrix}} \frac{P^\alpha}{[1 - e^{-t}(1-p)]^\alpha}$$