LG-1
SG-1

Unit 5:
Compound and mixed type distribution.

Introduction

A random variable $S = X_1 + X_2 + \dots + X_N$ is said to have a compound distribution if,

(i) N is a random variable

(ii) X_is are iid (identically, independently distributed)

(iii) Each X_is ~~are~~ is independent of N.

Mean and Variance of 'S'

$E(S) = E(X) \cdot E(N)$

$V(S) = E(N) V(X) + V(N) [E(X)]^2$

Q Suppose number of claims has a Binomial (100, 0.01) distribution and claim sizes have a Gamma (10, 0.02) distribution. Find the mean and variance of aggregate claim amount 'S'.

Solution

Let, N = no. of claims

X = claim amount

Since $N \sim B(100, 0.01)$ and

$X \sim G(10, 0.02)$

$\therefore E(N) = 100(0.01)$

$= 1$

$X \sim B(n,p)$

$\Rightarrow E(X) = np$

$V(X) = npq$

$X \sim G(\alpha, \beta)$

$E(X) = \alpha/\beta$

$V(X) = \alpha/\beta^2$

$\therefore V(N) = 100(0.01)(1-0.01)$

$= 0.99$

$E(X) = 10/0.02 = 500$

& $V(X) = \frac{10}{(0.02)^2} = \frac{500}{0.02} = 25000$

Now, the mean and variance of aggregate claim amount: S

$E(S) = E(N) \cdot E(X) \quad =$

$V(S) = E(N)\, V(X) + V(N)\,[E(X)]^2$

$\quad =$

H/W If X has an exponential distribution with parameter $\theta = 10$ and N has a Poisson with $\lambda = 50$ distribution, calculate the expected value and variance of 'S'.

$X \sim Exp(\theta)$

$E(X) = \dfrac{1}{\theta} \qquad V(X) = \dfrac{1}{\theta^2}$

$X \sim P(\lambda)$

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

classmate

Date

Page

Nov 16, 2024

[Compound Negative Exponential Distribution]

Poisson Distribution

Exponential Distⁿ

Let us assume that N follows Poisson distribution with parameter λ, and each X_i follows negative exponential distribution with parameter μ. The compound distribution ~~with~~ Y is the sum of N independent and identically distributed negative exponential random variables.

$$\text{Here, } N \sim P(\lambda) \Rightarrow P(N=n) = \frac{e^{-\lambda}\lambda^n}{n!}$$

$$\text{and } X_i \sim Exp(\mu) \Rightarrow f(x_i) = \mu e^{-\mu x_i}; \quad x_i \geq 0$$

Now, the compound distribution Y is given by the convolution of distⁿ of N and X_i's

$$f_Y(y) = \sum_{n=0}^{\infty} P(N=n)\, f_{x_1+x_2+\dots+x_n}(y)$$

Assuming independence, the sum of N negative exponential random variables follows a Gamma distribution with the shape parameter 'n' and rate parameter 'μ'.

$$\therefore f_{x_1+x_2+\dots+x_n}(y) = \frac{\mu^n y^{n-1} e^{-\mu y}}{(n-1)!} \qquad \begin{pmatrix} \alpha = n \\ \beta = \mu \end{pmatrix}$$

$\therefore$ The pdf of compound negative exponential distribution is given by,

$$f_Y(y) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \cdot \frac{\mu^n y^{n-1} e^{-\mu y}}{(n-1)!}$$

Mean and Variance:

Let N be a Poisson distribution with parameter λ representing the number of events and let $x_1, x_2, \ldots, x_N$ be independently and identically distributed random variables following exponential $dist^n$ with the parameter μ.

Then the mean and variance of compound negative exponential distribution is given by,

$$E(Y) = E(N)\,E(X) \quad [Y = x_1 + x_2 + \ldots + x_N]$$

$$= \lambda \cdot \frac{1}{\mu}$$

$$= \frac{\lambda}{\mu}$$

and $V(Y) = E(N)\,V(X) + V(N)\,[E(X)]^2$

$$= \lambda \frac{1}{\mu^2} + \lambda \left[\frac{1}{\mu}\right]^2$$

$$= \frac{\lambda}{\mu^2} + \frac{\lambda}{\mu^2}$$

$$\boxed{\therefore V(Y) = \frac{2\lambda}{\mu^2}}$$

[6M] Sure

Mixed distribution:

A mixed random variable is a type of random variable that follows a mixture distribution. A mixture distribution is a probability distribution derived from a combination of two or more component distributions.

$f(x)$ = prob. density function (pdf)

$F(x)$ = Cumulative distribution function (cdf).

classmate

Date

Page

Q. The probability distⁿ function of a random variable X is

$$f(x) = \begin{cases} x & \text{for} \quad 0 < x < 1 \\ 2-x & \text{for} \quad 1 \le x < 2 \\ 0 & \text{for} \quad x \ge 2 \end{cases}$$

Compute the cumulative distribution of X.

$= P(X \le x)$

$-\infty$ ② $+\infty$

$= \sum_{x} P(x)$ [Discrete

$= \int_{-\infty}^{x} f(x)\,dx$ [Continuous

$f(x) = \frac{dF(x)}{dx}$

Solution

For any x in the range $x \le 0$, we have:

$$F(x) = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{x} 0\,dx = 0$$

For x in the range $0 < x < 1$, we have

$$F(x) = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{x} f(x)\,dx$$

$$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{x} x\,dx$$

$$\boxed{F(x) = \frac{x^2}{2}}$$

For x in the range $1 \le x < 2$, we have

$$F(x) = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{x} f(x)\,dx$$

$$= 0 + \int_{0}^{1} x\,dx + \int_{1}^{x} (2-x)\,dx$$

$$= 0 + \frac{1}{2} + \left[2x - \frac{x^2}{2}\right]_1^x$$

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

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

$$F(x) = \left(2x - \frac{x^2}{2} - 1\right)$$

For x in the range $x \geq 2$

$$F(x) = \int_{-\infty}^{0} f(x)\,dx + \int_0^1 f(x)\,dx + \int_1^2 f(x)\,dx + \int_2^{\infty} f(x)\,dx$$

$$= 0 + \frac{1}{2} + \int_1^2 (2-x)\,dx + \int_2^{\infty} 0\,dx$$

$$= 0 + \frac{1}{2} + \left[2x - \frac{x^2}{2}\right]_1^2 + 0$$

$$= \frac{1}{2} + \left(4 - 2 - 2 + \frac{1}{2}\right)$$

$$= 1 + 0$$

$$= \boxed{1}$$

Thus, we have,

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x^2}{2} & 0 < x < 1 \\ 2x - \frac{x^2}{2} - 1 & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

classmate

Date
Page

H/W Find the distribution function of random variable X whose density function is given by

$$f(x) = \frac{x}{2} \quad \text{for } 0 < x < 1$$

$$= \frac{1}{2} \quad \text{for } 1 \le x < 2$$

$$= \frac{3-x}{2} \quad \text{for } 2 \le x < 3$$

$$= 0 \quad \text{for } x \ge 3$$

Q The time taken to perform a particular task, x hrs, has the probability density function

$$f(x) = \begin{cases} 10\,cx^2 & ; \ 0 \le x < 0.6 \\ 9c(1-x) & ; \ 0.6 < x < 1.0 \\ 0 & ; \ \text{otherwise} \end{cases}$$

where 'c' is a constant

(a) Find the value of 'c'
(b) Find the expected time
(c) Determine the prob. that the time will be
 (i) more than 0.8 hrs
 (ii) betn 0.4 and 0.8 ~~minute~~. hours

$$\int_x f(x)\,dx = 1$$

Solution

(a) $\because \int_x f(x)\,dx = 1$

$$\Rightarrow \int_0^1 f(x)\,dx = 1$$

$$\Rightarrow \int_0^{0.6} f(x)\,dx + \int_{0.6}^{1} f(x)\,dx = 1$$

$$\Rightarrow 10c\int_0^{0.6} x^2\,dx + 9c\int_{0.6}^{1} (1-x)\,dx = 1$$

$$\Rightarrow 10c\left[\frac{x^3}{3}\right]_0^{0.6} + 9c\left[x - \frac{x^2}{2}\right]_{0.6}^{1} = 1$$

$$\Rightarrow 10c\left(\frac{0.6^3}{3}\right) + 9c\left[1/2 - 0.6 + \frac{(0.6)^2}{2}\right] = 1$$

$$\Rightarrow 0.72c + 0.72c = 1$$

$$\therefore \boxed{c = \frac{1}{1.44}} = \underline{0.6944}$$

$$\therefore f(x) = \begin{cases} 10/1.44\, x^2; & 0 \le x < 0.6 \\ \frac{9}{1.44}(1-x); & 0.6 \le x < 1.0 \\ 0; & \text{otherwise} \end{cases}$$

(b)

$$E(X) = \int_x x\, f(x)\,dx$$

$$= \frac{10}{1.44}\int_0^{0.6} x^3\,dx + \frac{9}{1.44}\int_{0.6}^{1} x(1-x)\,dx$$

$$= \frac{10}{1.44}\left[\frac{x^4}{4}\right]_0^{0.6} + \frac{9}{1.44}\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_{0.6}^{1}$$

$$= \frac{10}{1.44}\left(\frac{0.6^4}{4}\right) + \frac{9}{1.44}\left[\frac{1}{2} - \frac{1}{3} - \frac{(0.6)^2}{2} + \frac{(0.6)^3}{3}\right]$$

$$= 0.225 + 6.25 * (-0.01 + 0.072)$$

$$= \cancel{0.225 + 0.3875}$$

$$= \cancel{0.612}$$

$$= \boxed{0.592}$$

classmate
Date
Page

(c) (i) $P(X > 0.8) = \int_{0.8}^{1} f(x)\, dx = \frac{9}{1.44} \int_{0.8}^{1} (1-x)\, dx$

$= \frac{9}{1.44} \left[x - \frac{x^2}{2}\right]_{0.8}^{1} = \frac{9}{1.44}\left[1 - \frac{1}{2} - 0.8 + \frac{0.8^2}{2}\right]$

$= \boxed{0.125}$

(ii) $P(0.4 < x < 0.8) = \int_{0.4}^{0.6} f(x)\, dx + \int_{0.6}^{0.8} f(x)\, dx$

$= \frac{10}{1.44} \int_{0.4}^{0.6} x^2\, dx + \frac{9}{1.44} \int_{0.6}^{0.8} (1-x)\, dx$

$= \frac{10}{1.44} \left[\frac{x^3}{3}\right]_{0.4}^{0.6} + \frac{9}{1.44} \left[x - \frac{x^2}{2}\right]_{0.6}^{0.8}$

$= \frac{10}{1.44} \left(\frac{0.6^3 - 0.4^3}{3}\right) + \frac{9}{1.44} \left[0.8 - \frac{(0.8)^2}{2} - 0.6 + \frac{(0.6)^2}{2}\right]$

$= 0.352 + 0.375$

$= \boxed{0.727}$

Q. Let, X have a mixed distribution with F(x) defined as follows:

$$F(x) = \begin{cases} 0 & \text{if } x < -2 \\ \frac{1}{12}(2+x) & \text{if } -2 \le x < 0 \\ \frac{1}{6}(1+x) & \text{if } 0 \le x < 4 \\ \frac{1}{12}(6+x) & \text{if } 4 \le x < 6 \\ 1 & \text{if } x \ge 6 \end{cases}$$

obtain the mean and variance of X.

CamScanner

Solution

The pdf does not exist for $x < -2$ and $x \geq 6$.

Since, $f(x) = \dfrac{dF(x)}{dx}$

For $-2 \leq x < 0$, $f(x) = \dfrac{d}{dx}\left(\frac{1}{12}(2+x)\right) = \frac{1}{12}$

$0 \leq x < 4$, $f(x) = \dfrac{d}{dx}\left(\frac{1}{6}(1+x)\right) = \frac{1}{6}$

$4 \leq x < 6$, $f(x) = \dfrac{d}{dx}\left(\frac{1}{12}(6+x)\right) = \frac{1}{12}$

Mean:

$$E(X) = \int_x x\, f(x)\, dx$$

$$= \int_{-2}^{0} x\, f(x)\, dx + \int_{0}^{4} x\, f(x)\, dx + \int_{4}^{6} x\, f(x)\, dx$$

$$= \int_{-2}^{0} \frac{1}{12} x\, dx + \int_{0}^{4} \frac{1}{6} x\, dx + \int_{4}^{6} \frac{1}{12} x\, dx$$

$$= \left[\frac{x^2}{24}\right]_{-2}^{0} + \left[\frac{x^2}{12}\right]_{0}^{4} + \left[\frac{x^2}{24}\right]_{4}^{6}$$

$$= -\frac{1}{6} + \frac{16}{12} + \frac{36}{24} - \frac{16}{24}$$

$$= \frac{-8 + 64 + 72 - 32}{48}$$

$$= \frac{104 - 8}{48} = \frac{96}{48} = \boxed{2}$$

classmate

Date

Page

Variance

$$V(X) = E(X^2) - [E(X)]^2$$

Now, $E(X^2) = \int_x x^2 f(x)\, dx$

$$= \int_{-2}^{0} x^2 f(x)\, dx + \int_{0}^{4} x^2 f(x)\, dx + \int_{4}^{6} x^2 f(x)\, dx$$

$$= \int_{-2}^{0} \frac{1}{12} x^2\, dx + \int_{0}^{4} \frac{1}{6} x^2\, dx + \int_{4}^{6} \frac{1}{12} x^2\, dx$$

$$= \left[\frac{x^3}{36}\right]_{-2}^{0} + \left[\frac{x^3}{18}\right]_{0}^{4} + \left[\frac{x^3}{36}\right]_{4}^{6}$$

$$= \frac{8}{36} + \frac{64}{18} + \frac{\cancel{196}\ 216}{36} - \frac{64}{36}$$

$$9.77 - 2^2 = \_\_\_$$

$$= \frac{8 + 128 + \cancel{196}\ 216 - 64}{36}$$

$$= 7.44$$

[check] $\therefore V(X) = 7.44 - (2)^2 = \_\_\_ - 4 = \_\_\_$

Extreme Value Distribution (EVD)

EVD is a statistical distribution used to model the extreme values (either maximum or minimum) in a set of observations. It is particularly relevant in fields of insurance, hydrology, finance and environmental science, where the focus is on ~~extreme science~~ extreme events.

If we are dealing with losses that have typical sizes, i.e. ones whose values come from the central part of the distⁿ, we can make use of central limit theorem. This tells us that, if we calculate the mean, $\bar{x}$ of a set of 'n' values ~~that~~ taken from a loss distribution ~~of~~ that has mean μ and variance σ^2, the standardized value, $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$, can be approximated using the standard normal distribution.

However, the most financial part of a loss distⁿ is usually the right-hand tail where the large losses occur. These are the extreme values of distⁿ. So, there is a similar way to approximate the behavior of the extreme values in the tail of the distⁿ.

The Generalized Extreme Value (GEV) distⁿ is a family of distⁿ with the EVD.

The pdf of GEV distⁿ is given by:

$$f(x) = \frac{1}{\beta}\left[1+\frac{\gamma(x-\alpha)}{\beta}\right]^{-(1+1/\gamma)} \exp\left[-\left(1+\frac{\gamma(x-\alpha)}{\beta}\right)^{1/\gamma}\right]$$

where, X = the random variable representing extreme values

γ = The shape parameter, determining the type of distribution

α = a location parameter
β = a scale parameter.

Applications of EVD

(i) EVD is widely used in hydrology to model the distribution of extreme values of river flows, rainfall and flood peaks. This is crucial for design infrastructure such as dams, bridges etc.

(ii) In finance, extreme value analysis is used for modeling extreme returns and assessing the risk associated with extreme market events.

(iii) In the insurance industry, extreme value distribution are used to model the tail risk associated with rare and severe events.
This is crucial setting insurance premiums
(iii) and reserves to cover potential extreme losses.

(iv) In the field of telecommunications, extreme value analysis is applied to model extreme traffic loads on communication network. This helps in designing network. This helps in designing network that can handle rare but intense periods of high demand.