Questionn The time between failures of a laser machine is exponentially distributed with a mean of 25,000 hours. What is the expected time until the second failure? What is the probability that the time until the third failure exceeds 50,000 hours?

2.2.4 Beta Distribution

Beta Function

We define a beat function as $\mathrm{B}(\alpha,\beta)$ as $\mathrm{B}(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$ with parameters $\alpha>0$ and $\beta>0$

Theorem:

i) $\mathrm{B}(\alpha,\beta)=\mathrm{B}(\beta,\alpha)$ ie by putting $y=1-x$ in the function

ii) $\mathrm{B}(\alpha,\beta)=\int_0^\infty \frac{u^{\alpha-1}}{(1+u)^{\beta-1}}du$ ie by putting $u=\frac{x}{1-x}$ or $x=\frac{u}{1+u}$ in the function

iii) $\mathrm{B}(\alpha,\beta)=2\int_0^{\pi/2}\cos^{2\alpha-1}t\sin^{2\beta-1}t\,dt$ ie by putting $x=\cos^2 t$ in the function

iv) $\mathrm{B}(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

v) $\Gamma(\tfrac{1}{2})=\sqrt{\pi}$ use $\mathrm{B}(\tfrac{1}{2},\tfrac{1}{2})=\frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(1)}=[\Gamma(\tfrac{1}{2})]^2$ but

$\mathrm{B}(\tfrac{1}{2},\tfrac{1}{2})=2\int_0^{\pi/2}\cos^{2\times\frac{1}{2}-1}t\sin^{2\times\frac{1}{2}-1}t\,dt=2\int_0^{\pi/2}dt=\pi \quad \Rightarrow \quad \Gamma(\tfrac{1}{2})=\sqrt{\pi}$

Note $\Gamma(\tfrac{1}{2})=\int_0^\infty x^{-1/2}e^{-x}dx=2\int_0^\infty e^{-u^2}du=\sqrt{\pi}$. Simply $\int_0^\infty e^{-x^2}dx=\tfrac{1}{2}\sqrt{\pi}$

Beta Distribution

Definition: A random variable X is said to have a standard beta distribution with parameters α and β if it's probability density function is given by

$$f(x)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)},0\le x\le 1 \text{ and } f(x)=0 \textit{ elsewhere} \text{ we denote this as } X\sim Beta(\alpha,\beta)$$

Theorem: If X has a standard beta distribution with parameters α and β, then

$$E(X)=\mu=\frac{\alpha}{\alpha+\beta} \quad \text{and.} \quad Var(X)=\sigma^2=\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

3. MOMENTS AND MOMENT-GENERATING FUNCTIONS

Definition: The k[th] **moment** of a r.v. X taken about zero, or about origin is defined to be $E[X^k]$ and denoted by μ_k' .

Definition: The k[th] **moment** of a r.v. X taken about its mean, or the k[th] **central moment** of X, is defined to be $E(X-\mu)^k$ and denoted by μ_k .

Definition: The **moment-generating function** (mgf), $m(t)$, for a r.v. X is defined to be $M_x(t)$ or simply $M(t)=E[e^{tx}]$

We say that an mgf for X exists if there is $b>0$ such that $M(t)<\infty$ for $|t|\le b$.

$$e^{tx}=1+\frac{tx}{1!}+\frac{(tx)^2}{2!}+\frac{(tx)^3}{3!}+\frac{(tx)^4}{4!}+.....$$

$$M(t)=E\left[e^{tx}\right]=\sum_{all\,x}\left(1+\frac{tx}{1!}+\frac{(tx)^2}{2!}+\frac{(tx)^3}{3!}+\frac{(tx)^4}{4!}+....\right)P(X=x)$$

$$=\sum_{all\,x}P(X=x)+t\sum_{all\,x}xP(X=x)+\frac{t^2}{2!}\sum_{all\,x}x^2P(X=x)+\frac{t^3}{3!}\sum_{all\,x}x^3P(X=x)+....$$

$$=1+tE(X)+\frac{t^2}{2!}E(X^2)+\frac{t^3}{3!}E(X^3)+\ldots \text{ie a function of all moments about the origin}$$

Theorem: If $M(t)$ exists, then for any $k\in N$ $\left.\frac{d^kM(t)}{dt^k}\right|_{t-0}=\mu_k^{/}=E(X^k)$

Proof

$$M(t)=1+t\mu_1^{/}+\frac{t^2}{2!}\mu_2^{/}+\frac{t^3}{3!}\mu_3^{/}+\ldots$$

$$M^{|}(t)=\mu_1^{/}+\frac{2t}{2!}\mu_2^{/}+\frac{3t^2}{3!}\mu_3^{/}+\ldots\quad\Rightarrow\quad M^{|}(0)=\mu_1^{/}=E(X)$$

$$M^{||}(t)=\mu_2^{/}+\frac{2t}{2!}\mu_3^{/}+\frac{3t^2}{3!}\mu_4^{/}+\ldots\quad\Rightarrow\quad M^{||}(0)=\mu_2^{/}=E(X^2)$$

Remark: The mgf of a particular distribution is unique and we can recognize the pdf if we are given the mgf.

Example 1

The mgf of a r.v Y is given by $M(t)=\frac{1}{6}e^t+\frac{1}{3}e^{2t}+\frac{1}{2}e^{3t}$ Find the mean and variance of Y

Solution

$$E(Y)=M^{|}(0)=\left.\left(\tfrac{1}{6}e^t+\tfrac{2}{3}e^{2t}+\tfrac{3}{2}e^{3t}\right)\right|_{t-0}=\tfrac{1}{6}+\tfrac{2}{3}+\tfrac{3}{2}=\tfrac{7}{3}$$

$$E(Y^2)=M^{||}(0)=\left.\left(\tfrac{1}{6}e^t+\tfrac{4}{3}e^{2t}+\tfrac{9}{2}e^{3t}\right)\right|_{t-0}=\tfrac{1}{6}+\tfrac{4}{3}+\tfrac{9}{2}=6$$

$$Var(Y)=E(Y^2)-\mu^2=6-\left(\tfrac{7}{3}\right)^2=\tfrac{5}{9}$$

Example 2 Find the mgf of a r.v $\mathrm{X\sim Po}(\lambda)$

Solution

$$M(t)=E\left[e^{tx}\right]=\sum_{x=0}^{\infty}\left[e^{tx}\right]\frac{\lambda^x e^{-\lambda}}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(\lambda e^t\right)^x}{x!}=e^{-\lambda}e^{\lambda e^t}=e^{\lambda\left(e^t-1\right)}$$

Now $E(X)=M^{|}(0)=\left.\lambda e^t e^{\lambda\left(e^t-1\right)}\right|_{t-0}=\lambda$

Example 3 Find the mgf of a r.v X whose pmf is given by $\mathrm{f(x)}=\begin{cases}\frac{1}{6}\left(\frac{5}{6}\right)^x \text{ for } x=0,1,2,3.....\\ 0\quad elsewhere\end{cases}$

hence obtain the mean and variance of X

Solution

$$.\,M(t)=E\left[e^{tx}\right]=\sum_{x=0}^{\infty}e^{tx}\,\tfrac{1}{6}\left(\tfrac{5}{6}\right)^x=\sum_{x=0}^{\infty}\tfrac{1}{6}\left(\tfrac{5}{6}e^t\right)^x=\frac{\frac{1}{6}}{1-\frac{5}{6}e^t}=\frac{1}{6-5e^t}=\left(6-5e^t\right)^{-1}$$

$M^{|}(t) = 5e^t\left(6-5e^t\right)^{-2}$ and $M^{||}(t) = 5e^t\left(6-5e^t\right)^{-2} + 50e^{2t}\left(6-5e^t\right)^{-3}$

$\Rightarrow E(X) = 5e^t\left(6-5e^t\right)^{-2}\Big|_{t-0} = 5$ and $E(Y^2) = \left[5e^t\left(6-5e^t\right)^{-2} + 50e^{2t}\left(6-5e^t\right)^{-3}\right]_{=0} = 55$

$Var(X) = E(X^2) - \mu^2 = 55 - 5^2 = 30$

Exercise

1) The mgf of a r.v Y is given by; a) $M(t) = e^{2t^2+3t}$ b) $M(t) = \exp\left\{\frac{1}{2}\sigma^2 t^2 + t\mu\right\}$ Find the mean and variance of Y
2) A r.v X has a gamma distribution with parameters , Find the mgf of X hence obtain the mean and variance of X

3.1 The Mgf of a Sum of Independent Random Variables

The mgf of the sum of n independent random variable is the product of their individual mgf's
The mean (variance) of the sum of n independent random variable is the sum of their individual means (variances).
The mgf about $X = a$ is given by $M_{x,a}(t) = E\left[e^{t(x-a)}\right] = e^{-at}E\left[e^{tx}\right] = e^{-at}M_x(t)$

4 NORMAL DISTRIBUTION

4.1 Introduction

The normal, or Gaussian, distribution is one of the most important distributions in probability theory. It is widely used in statistical inference. One reason for this is that sums of random variables often approximately follow a normal distribution.

Definition A r.v X has a normal distribution with parameters μ and σ^2, abbreviated $X \sim N\left(\mu, \sigma^2\right)$ if it has probability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \text{ for } -\infty < x < \infty \text{ and } \sigma > 0$$

Where μ is the mean and σ is the standard deviation.

4.1.1 Properties of normal distribution

1) The normal distribution curve is bell-shaped and symmetric, about the mean
2) The curve is asymptotic to the horizontal axis at the extremes.
3) The highest point on the normal curve is at the mean, which is also the median and mode.
4) The mean can be any numerical value: negative, zero, or positive
5) The standard deviation determines the width of the curve: larger values result in wider, flatter curves
6) Probabilities for the normal random variable are given by areas under the curve. The total area under the curve is 1 (0.5 to the left of the mean and 0.5 to the right).
7) It has inflection points at $\mu-\sigma$ and $\mu+\sigma$.
8) Empirical Rule:
   a) 68.26% of values of a normal random variable are within ± 1 standard deviation of its mean. ie $P\left(\mu-\sigma \le X \le \mu+\sigma\right) = 0.6826$
   b) 95.44% of values of a normal random variable are within ± 2 standard deviation of its mean. ie $P\left(\mu-2\sigma \le X \le \mu+2\sigma\right) = 0.9544$
   c) 99.72% of values of a normal random variable are within ± 3 standard deviation of its mean. ie $P\left(\mu-3\sigma \le X \le \mu+3\sigma\right) = 0.9972$