Frequently Asked Questions

1. How do you find the probabilities of normal distribution?

Answer:

Usually we find the probabilities by integrating its probability density function. But is not a easy job to integrate the probability density function of normal distribution. We have Standard normal tables which will give the probabilities corresponding to the different ranges of standard normal variate. Hence, always we convert the normal variable into standard normal variable and using standard normal tables, we obtain the required probabilities.

2. Write mgf of normal distribution with parameters $\mu=10$ and $\sigma^2=9$.

Answer:

The mgf of normal distribution with parameters $\mu=10$ and $\sigma^2=9$ is given by,
$MX(t)=\exp(t\mu+t^2\sigma^2/2)=\exp(10t+4.5t^2)$

3. For a certain normal distribution, the first moment about 10 is 40 and the fourth moment about 50 is 48. What is the Expectation and Standard deviation of the distribution?

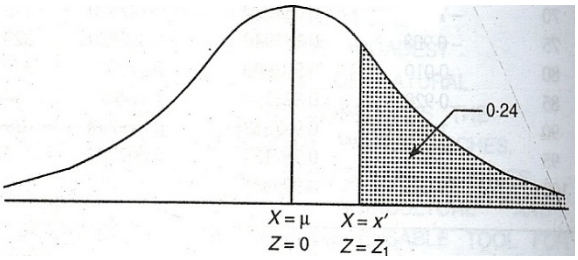
Answer:

We know that mu 1 dash is the first moment about the point X=A, the Expectation, the mean is given by, $A+\mu_1'$.
We are given μ_1'(about point X=10)=40
$\Rightarrow$Mean=10+40=50.
Also μ_4'(about the point X=50)=48
Since mean is also 50, $\mu_4'=\mu_4=48$
But for a normal distribution with standard deviation σ,
$\mu_4=3\sigma^4=48\Rightarrow \sigma^4=16$ or $\sigma=2$.

4. X is normally distributed and the mean of X is 12 and standard deviation is 4. Find probability of the following:
   i. (a) $X\geq 20$ (b)$X\leq 5$ (c)$0\leq X\leq 12$
   ii. Find x when $P(X>x)=0.24$
   iii. Find x_1 and x_2 when $P(x_1<X<x_2)=0.5$ and $P(X>x_1)=0.25$.

Answer:

We have$\mu=12$ and $\sigma=4$.
i.e. $X\sim N(12,16)$ and the corresponding standard normal variate is $Z={}^{(X-\mu)}/_{\sigma}={}^{(X-12)}/_{4}$
(i) We can find all the probabilities from standard normal tables. Hence, every time we convert the original variable X into standard normal variable Z.
a. $P(X\geq 20)=P[Z\geq (20-12)/4]=P(Z\geq 2)=0.0228$
b. $P(Z\leq 5)=P[Z\leq (5-12)/4]=P(Z\leq -1.75)=0.04006$
c. $P(0\leq X\leq 12)=P[(0-12)/4\leq Z\leq (12-12)/4]$
$=P(-3\leq Z\leq 0)=0.49865$
(ii) Given $P(X>x)=0.24\Rightarrow P(Z>z)=0.24$, where $z={}^{x-12}/_{4}$



From the above curve we have z=0.71

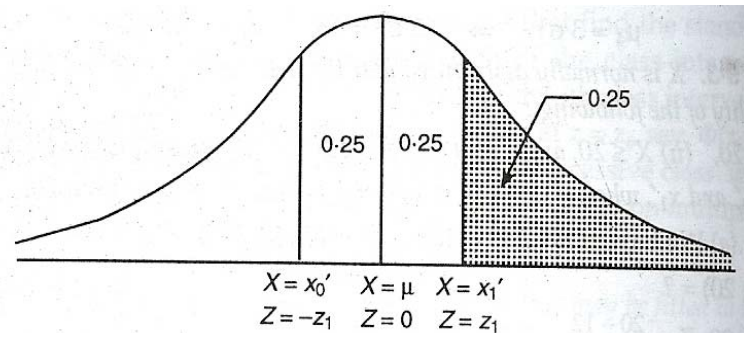
i.e., $^{x-12}/_4=0.71 \Rightarrow x=12+(4)(0.71)=14.84$

iii. We are given

$P(x_1<X<x_2)=0.5$ and $P(X>x_1)=0.25$.

i.e. $P(z_1<Z<z_2)=0.5$ and $P(Z>z_1)=0.25$, where $z_1={}^{x1-12}/_4$ and $z_2={}^{x2-12}/_4$

The points z1 and z2 are located as shown in the following curve.



From the curve

$P(Z>z_1)=P(Z<z_2)$ or $P(z<-z_1)=0.25$, as we know that normal curve is symmetric about zero. Hence from standard normal tables,

$z_1=0.67$ and $z_2=-z_1=-0.67$ or

$X_1=12+(4)(0.67)=14.68$ and $X_2=12+(4)(-0.67) = 9.32$

5. In a distribution exactly normal, 10.03% of the items are under 25 kilogram weight and 89.97% of the items are under 70 kilogram weight. What are the mean and standard deviation of the distribution?

Answer:

Let X denotes the weight of the items in kg. If $X \sim N(\mu, \sigma^2)$, then we are given $P(X<25)=0.1003$ and $P(X<70)=0.8997$

The points X=25 and X=70 are located as below.

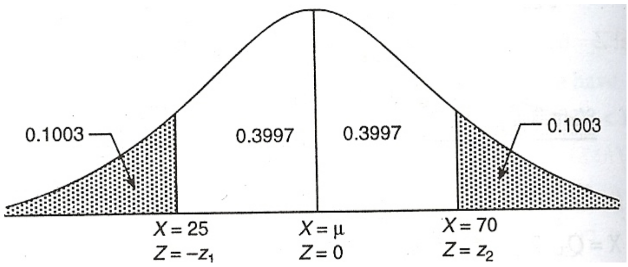
Since the value X=25 is located to left of the ordinate X=μ (because left to this value probability is<0.5), the corresponding value of Z is negative.

$Z={}^{(25-\mu)}/_{\sigma}=-z_1$--------------(1)

Similarly the value X=70 is located to right of the ordinate X=μ (because left to this value probability is > 0.5), the corresponding value of Z is positive)

$Z={}^{(70-\mu)}/_{\sigma}=z_2$--------------(2)

Let us draw the diagram for indicating the probabilities.



From the diagram, it is obvious that,

$P(Z<-z_1)=0.1003$ and $P(Z<z_2)=0.8997$

Using standard normal tables we get,

$z_2=1.28$ and $z_1=1.28$

Substituting the values of z_1 and z_2 in (1) and (2), we get,

$^{(25-\mu)}/_{\sigma}=-1.28 \Rightarrow 25-\mu=-1.28\sigma$--------------(3)

and$^{(70-\mu)}/_{\sigma}=1.28 \Rightarrow 70-\mu=1.28\sigma$ ---------(4)

Subtracting 3 from 4, we get, $45=2.56\sigma \Rightarrow \sigma=17.578$

Substituting the value of σ in (3) we have 25-μ=-1.28(3)⇒μ=47.5
Hence, the mean is 47.5 kilogram and standard deviation is 17.578 kilogram.

6. The daily consumption of milk in a city, in excess of 20,000 litres, is approximately distributed as gamma variate with parameters α=1/10000 and β=2. The city has a daily stock of 30,000 litres. What is the probability that the stock is insufficient on a particular day?

Answer:

If the random variable X denotes the daily consumption of milk (in litres) in a city, then the random variable Y=X-2000 has gamma distribution with pdf

$$g(y) = \frac{1}{(10000)^2\Gamma 2} y^{2-1} e^{-y/10000}; y > 0$$

Since the daily stock of the city is 30,000 litres, the required probability 'p' that the stock is insufficient on a particular day is given by,

$$p = P(X > 30000) = P(Y > 10000)$$

$$= \int_{10000}^{\infty} g(y)dy = \int_{10000}^{\infty} \frac{y^{2-1}e^{-y/10000}}{10000^2} dy = \int_1^{\infty} ze^{-z}dz, [z = y/10000]$$

$$= \left|-ze^{-z}\right|_1^{\infty} + \int_1^{\infty} e^{-z}dz = e^{-1} - \left|e^{-z}\right|_1^{\infty} = e^{-1} + e^{-1} = 2.e^{-1}$$

7. Write the mgf of exponential distribution with mean 0.5.

Answer:

In an exponential distribution, we know that mean is reciprocal of the parameter.
Hence parameter θ=1/0.5=2
Its mgf is given by, $(1\text{-}t/\theta)^{-1}=(1\text{-}t/2)^{-1}$.

8. Suppose that during rainy season on a tropical island the length of the shower has an exponential distribution with parameter 2, time being measured in minutes. What is the probability that a shower will last more than 3 minutes? If a shower has already lasted for 2 minutes, what is the probability that it will last or at least one more minute?

Answer:

Let X denote length of shower in minutes. Hence X~exp(θ),θ=2
Hence pdf is given by, $f(x)=\theta e^{-x\theta}=2e^{-2x}$, x>0
P[shower will last more than 3 minutes]

$$=P[X>3]= \int_3^{\infty} f(x)dx = 2\int_3^{\infty} e^{-2x}dx = 2\left|\frac{e^{-2x}}{-2}\right|_3^{\infty} = e^{-6}$$

P[shower which has already lasted for 2 minutes will last for at least one more minute]
Because of memory less property of exponential distribution, it is same as
P[shower will last for at least one more minute]

$$=P[x>1] = \int_1^{\infty} f(x)dx = 2\int_1^{\infty} e^{-2x}dx = 2\left|\frac{e^{-2x}}{-2}\right|_1^{\infty} = e^{-2}$$

9. We have given the mgf of a distribution is given by $1/(1\text{-}2t)^4$. Identify the distribution and hence find its mean and variance.

Answer:

By uniqueness theorem of mgf, we can always identify the distribution of a random variable.
We have given mgf
$M_X(t)= 1/(1\text{-}2t)^4$

It is similar to mgf of gamma distribution, which is of the form $\dfrac{1}{(1-t/\beta)^{\alpha}}$

By comparison, it is clear that given random variable has gamma distribution with two parameters α=5 and β=1/2=0.5.
Also we know that mean of gamma distribution is α/β=5/0.5=10
And variance = α/β^2 = $5/0.5^2$=20

10. If X is uniformly distributed with mean 1 and variance 4/3, find P(X<0).

Answer:

Let X~U(a,b) so that f(x)=1/(b-a), a<x<b
Mean (b+a)/2=1, Implies, b+a=2
Variance =$(b-a)^2/12=4/3$, Implies $(b-a)^2=16$ or (b-a)=±4
On solving we get, a=-1 and b=3 or a=3 and b=-1
But in uniform distribution, we should have, a<b, the solution for a and b is,
a=-1 and b=3
Therefore f(x)=¼;-1<x<3 .

Hence we can find $P(X<0)=\int_{-1}^{0} f(x)dx=\frac{1}{4}\left|X\right|_{-1}^{0}=\frac{1}{4}$

11. Suppose that the income of a certain population has the Pareto distribution with shape parameter 3 and scale parameter 1000. Find the proportion of the population with incomes between 2000 and 4000.

Answer:

Let X denotes the Income of a certain population. Hence X has Pareto distribution with shape parameter(α)=3 and scale parameter (b)=1000.
Hence we can write pdf as follows

$$f(x)=\frac{\alpha b^{\alpha}}{x^{\alpha+1}}=\frac{3\times 1000^3}{x^4};x\geq 1000$$

P(income between 2000 and 4000)
=P(2000<X<4000)

$$=\int_{2000}^{4000} f(x)dx=3\times 1000^3\int_{2000}^{4000} x^{-4}dx=3\times 1000^3\left|\frac{x^{-3}}{-3}\right|_{2000}^{4000}=1000^3(2000^{-3}-4000^{-3})=0.1094$$

12. Define log-normal distribution.

Answer:

If Y~N(μ, σ^2), then X=e^Y is called a log normal random variable

13. If Y~N(5,16), then define the distribution of X=eY. Obtain mean and variance. Also comment on skewness and kurtosis of the distribution of X.

Answer:

We know that if Y ~N(μ, σ^2), then X=e^Y is called a log normal random variable. r^{th} raw moment is given by, $\mu_r'=e^{r\mu+r^2\sigma^2/2}=e^{5r+16r^2/2}$

First raw moment, $\mu_1'=e^{5+16/2}=e^{13}$

2^{nd} raw moment is given by, $\mu_2'=e^{2\mu+2\sigma^2}=e^{2(5)+2(16)}=e^{42}$

Variance of the distribution is given by,
$\mu_2=\mu_2'-\mu_1'^2=e^{42}-(e^{13})^2=e^{26}(e^{16}-1)$
We know that the coefficient of skewness

$$\beta_1=\frac{\left[e^{3\sigma^2}-3e^{\sigma^2}+2\right]^2}{\left[e^{\sigma^2}-1\right]^3}=\frac{\left[e^{3(16)}-3e^{16}+2\right]^2}{\left[e^{16}-1\right]^3}>0$$

Also coefficient of kurtosis is given by,

$$\beta_2 = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 3 = e^{64} + 2e^{48} + 3e^{32} - 3 = 6.2351 \times 10^{27} > 3$$

Hence X is positively skewed and leptokurtic distribution.

14. Subway trains on a certain line run every half hour between mid-night and six in the morning. What is the probability that a man entering the station at a random time during this period will have to wait at least twenty minutes?

Answer:

Let the random variable X denotes the waiting time (in minutes) for the next train. Under the assumption that a man arrives at the station at random, X is distributed uniformly on (0, 30) with pdf f(x)=1/30; 0<x<30.

The probability that he has to wait at least 20 minutes is given by,

$$P(X \geq 20) = \int_{20}^{30} f(x)dx = \frac{1}{30}\int_{20}^{30} 1.dx = \frac{1}{30}\Big|X\Big|_{20}^{30} = \frac{1}{3}$$

15. Suppose that the amount of time one spends in a bank is exponentially distributed with mean 10 minutes, what is the probability that a customer will spend more than 15 minutes in the bank? What is the probability that a customer will spend more than 15 minutes in the bank given that he is still in the bank after 10 minutes?

Answer:

Let X denote amount of time one spends in a bank. Hence X~exp(θ)

Mean is 10 minutes. Hence θ=1/10=0.1

Therefore pdf is given by, $f(x)=\theta e^{-x\theta}$, x>0.

By substituting θ, we get $f(x)=(0.1)e^{-0.1x}$, x>0

Consider

P(Customer will spend more than 15 minutes)

$=P(X > 15) = 0.1\int_{15}^{\infty} e^{-0.1x}dx = e^{-0.1\times 15} = 0.22$

Now we need to find the probability that customer will spend more than 15 minutes in the bank given he is still in the bank after 10 minutes

i.e. P(X > 15|X > 10)

Using memory less property we can write,

$P(X > 15|X > 10) = P(X > 5) = e^{-0.1x5}= 0.604$