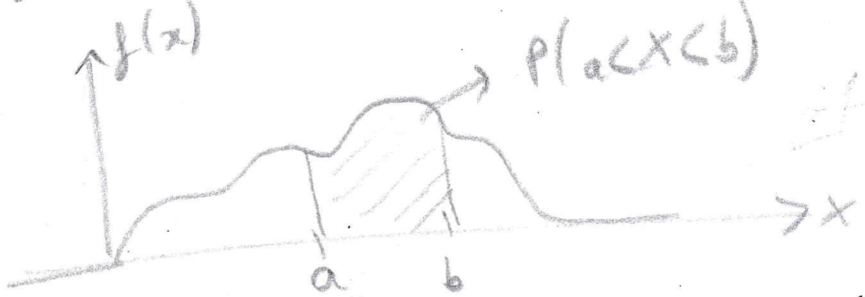
Continuous Rvs and Their Related Distributions

A continuous r.v. is a r.v. with an interval (either finite or infinite) of real numbers for its range.

A probability density function $f(x)$ can be used to describe the probability distribution of a continuous r.v. X.

If an interval is likely to contain a value for X, its prob. is large and it corresponds to a large values for $f(x)$

The prob. that X is between a and b is determined as the integral of $f(x)$ from a to b.



$f(x)$ is used to calculate an area that represents the prob. that X assumes a value in $[a,b]$.

For a continuous r.v. X, a prob. density function is a function such that

1) $f(x) \geq 0$

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

3) $P(a \leq X \leq b) = \int_a^b f(x)\,dx \Rightarrow$ area under $f(x)$ from a to b, for any a and b.

$0 \leq P(a < X < b) \leq 1$

If X is a continuous r.v. for any x_1 and x_2

$$P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2)$$

Ex Suppose that $f(x) = e^{-x}$ $x > 0$.

a) $P(X>1) = ?$

$$P(X>1) = \int_1^{\infty} e^{-x}\,dx = \frac{e^{-x}}{-1}\Big|_1^{\infty} = \frac{e^{-\infty}}{-1} + \frac{e^{-1}}{1}$$

$$= \frac{1}{e} = 0.3679$$

b) $$P(1<X<2.5) = \int_1^{2.5} e^{-x}\,dx = -e^{-x}\Big|_1^{2.5} = -e^{-2.5} + e^{-1}$$

$$= e^{-1} - e^{-2.5}$$

$$= \frac{1}{e} - \frac{1}{e^{2.5}}$$

$$= 0.2858$$

c) $P(X=3) = 0$

d) $$P(X<4) = \int_0^{4} e^{-x}\,dx = \frac{e^{-x}}{-1}\Big|_0^{4} = -e^{-4} + e^{-0} = -(e^{-4} - e^{-0})$$

$$= 1 - e^{-4} = 0.9817$$

e) $$P(X \leq 3) = \int_0^{3} e^{-x}\,dx = \frac{e^{-x}}{-1}\Big|_0^{3} = -e^{-3} + e^{-0}$$

$$= 1 - e^{-3} = 0.0498$$

Cumulative Dist. Function of a Continuous r.v

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du \quad \text{for } -\infty < x < \infty$$

The prob. density function of a continuous r.v. can be determined from the cum. dist. func. by differentiating

$$\frac{\partial}{\partial x}\int_{-\infty}^{x} f(u)\,du = f(x)$$

<u>Ex</u> Suppose the c.d.f of a r.v. X is

$$F(x) = \begin{cases} 0, & x<0 \\ 0.2x, & 0\le x<5 \\ 1, & 5\le x \end{cases} \longrightarrow f(x) = \begin{cases} 0.2 & 0\le x<5 \\ 0, & \text{o.w.} \end{cases}$$

a) $P(X<2.8)$

$F(2.8) = 0.2\times 2.8 = 0.56 \longrightarrow \int_0^{2.8} 0.2\,dx = 0.2x\Big|_0^{2.8} = 0.2\times 2.8$ ✓

b) $P(X>1.5) \longrightarrow \int_{1.5}^{5} 0.2\,dx = 0.2x\Big|_{1.5}^{5} = 1-0.3 = 0.7$

$P(X<1.5) = F(1.5) = 0.2\times 1.5 = 0.3$

30

$1-0.3 = 0.7$

c) $P(X<-2) = 0$

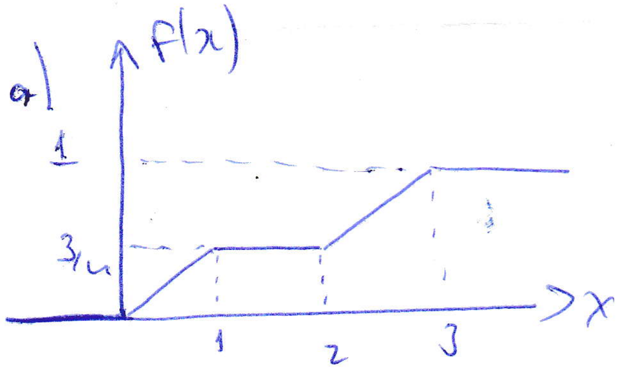
d) $P(X>6) = 1$

Ex We have the c.d.f. of a r.v. X as

$$F_X(x) = \begin{cases} 0, & x<0 \\ \frac{3x}{4}, & 0 \le x < 1 \\ \frac{3}{4}, & 1 \le x < 2 \\ \frac{x}{4}+\frac{1}{4}, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

a) Sketch $F_X(x)$.
b) Find $f(x)$.
c) find $P(1 < X < 2)$
d) Find $P(2 \le X \le 3)$

a)



b) for $x<0$, $f(x)=0$

for $0 \le x < 1$, $f(x) = \frac{\partial}{\partial x}\frac{3x}{4} = \frac{3}{4}$

for $1 \le x < 2$, $f(x) = \frac{\partial}{\partial x}\left(\frac{3}{4}\right) = 0$

for $2 \le x < 3$, $f(x) = \frac{\partial}{\partial x}\left(\frac{x}{4}+\frac{1}{4}\right) = \frac{1}{4}$

for $x \ge 3$, $f(x) = \frac{\partial}{\partial x}(1) = 0$.

$$f_X(x) = \begin{cases} \frac{3}{4}, & 0 \le x < 1 \\ \frac{1}{4}, & 2 \le x < 3 \\ 0, & \text{o.w.} \end{cases}$$

Let's check if we can have $F_X(x)$ from $f_X(x)$.

for $0 \le x < 1$, $F(x) = \int_0^x \frac{3}{4}\,du = \frac{3}{4}x$

for $1 \le x < 2$, $F(x) = \int_0^1 \frac{3}{4}\,du + \int_1^x 0\,du = \frac{3}{4}$

for $2 \le x < 3$, $F(x) = \int_0^1 \frac{3}{4}\,dx + \int_1^2 0\,dx + \int_2^x \frac{1}{4}\,dx = \frac{3}{4} + \left(\frac{x}{4} - \frac{2}{4}\right)$

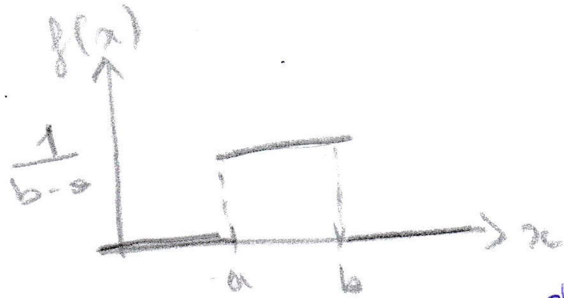
$= \frac{1}{4} + \frac{x}{4}$

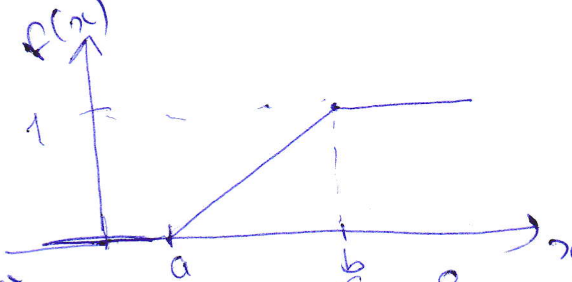
Continuous Uniform Distribution

A continuous r.v. X with prob density function,

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{o.w.} \end{cases}$$

is a continuous uniform r.v.



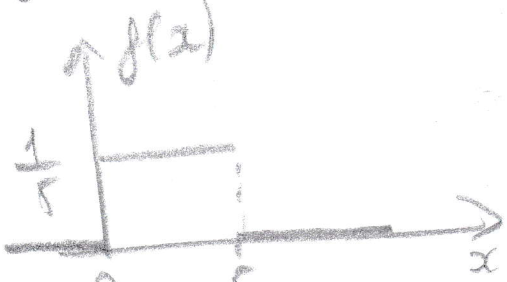


$$F(x) = \int_a^x \frac{1}{b-a}\,dx \rightarrow F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \leq x < b \\ 1, & x \geq b \end{cases}$$

Ex [illegible]

Ex Suppose that I take a bus to work and that every five minutes a bus arrives at my stop. Because of variation in time that I leave my house, my waiting time X is a continuous r.v. varying between [0,5]. The p.d.f of X will be;

$$f(x) = \begin{cases} \frac{1}{5}, & 0 \leq x \leq 5 \\ 0, & \text{o.w.} \end{cases}$$



which is a uniform dist. between 0 and 5.

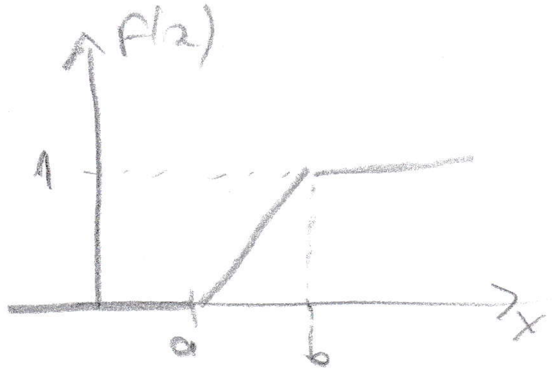
What is the prob. of waiting more than 4 minutes?

$$P(X>4) = \int_4^5 f(x)\,dx = \int_4^5 \frac{1}{5}\,dx = \frac{5}{5} - \frac{4}{5} = \frac{1}{5}$$

Ex Obtain the c.d.f of uniform dist.

$$F(x) = P(X \leq x) = \int_a^x \frac{1}{b-a}\,du = \frac{1}{b-a}\, u \Big|_a^x = \frac{x-a}{b-a}, \text{ then}$$

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



* NORMAL DISTRIBUTION

The most widely used dist.

A r.v. X with p.d.f

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty$$

is a normal r.v. with parameters μ, where $-\infty < \mu < \infty$, and $\sigma > 0$.

$$X \sim N(\mu, \sigma^2)$$

Ex. 4-10 acikla

Some useful results concerning a normal dist:

$$P(\mu - \sigma < X \leq \mu + \sigma) = 0.6827$$
$$P(\mu - 2\sigma < X \leq \mu + 2\sigma) = 0.9545$$
$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.9973$$

Also, from the symmetry of $f(x)$, $P(X>\mu) = P(X<\mu) = 0.5$

Prob. density funct ↓ as x moves farther from μ.

Consequently, the prob. that a measurement falls far from μ is small, and at the same time at some distance from μ, the prob. of an interval can be approximated as zero.

$6\sigma \to$ referred to as a width of normal dist. because more than 0.9973 of the prob. of a normal dist. is within the interval $(\mu-3\sigma, \mu+3\sigma)$

The area under the normal prob. den. func. from $-\infty < x < \infty$ is 1.

Standart Normal Random Variable

A normal r.v. with

$$\mu = 0 \text{ and } \sigma^2 = 1,$$

is called a standart normal r.v. and is denoted by Z.

The c.d.f of a standart normal r.v. is denoted as

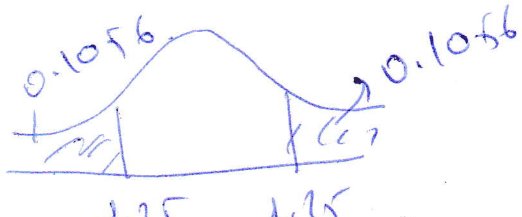
$$\Phi(z) = P(Z \leq z)$$

Appendix Table III provides cumulative probabilities for a standart normal r.v.

Ex a) $P(Z \leq 1.25) = 0.8944$

b) $P(Z > 1.25) = 1 - P(Z \leq 1.25) = 1 - 0.8944$
$= 0.1056$

c) $P(Z \leq -1.25) = 0.1056$



d) $P(-0.38 \leq Z \leq 1.25) = P(Z \leq 1.25) - P(Z < -0.38)$
$= 0.8944 - 0.3520$
$= 0.5424$

Ex Find z which satisfies $P(Z \leq z) = 0.8944$
$\downarrow$
1.25

$P(Z \leq z) = 0.8945$
$\downarrow$
2.54.

Standardizing a normal r.v.

If $X \sim N(\mu, \sigma^2)$, the r.v. $Z = \frac{X - \mu}{\sigma}$

is a normal r.v with mean 0 and variance 1. That is, Z is standart normal r.v.

The r.v. Z represents the distance of X from its mean in terms of standard deviations.

Suppose X is a normal r.v. with mean μ and variance σ^2.

Then,

$$P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = P(Z \le z)$$

where Z is the standard normal r.v. and $z = \frac{(x-\mu)}{\sigma}$ is the z-value obtained by standardizing X. The prob. is obtained by using Appendix Table III with $z = \frac{x-\mu}{\sigma}$.

Ex Let $X \sim N(\mu = 10, \sigma^2 = 25)$

Then find $P(X \le 20) = ?$

$$P\Big(\underbrace{\frac{X-\mu}{\sigma}}_{Z} \le \frac{20-10}{5}\Big) \Rightarrow P(Z \le 2) = 0.9772 \text{ (from table)}$$

Let $X_1, X_2, \ldots, X_{25}$ be independent and identically distributed r.v.s. each have normal dist with mean $\mu = 20$ and variance $\sigma^2 = 16$. Then,

$P(\bar{X} > 22) = ?$

If $X \sim N(\mu, \sigma^2) \Rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ (Central Limit Theorem)

$$P(\bar{X} > 22) = P\Big(\underbrace{\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}_{Z} > \frac{22-20}{\frac{4}{\sqrt{25}}}\Big) = P\Big(Z > \frac{2}{\frac{4}{5}}\Big)$$

$$= P\Big(Z > \frac{10}{4}\Big) = P(Z > 2.5) = 1 - P(Z \le 2.5)$$

$$= 1 - 0.9938$$

$$= 0.0062$$

EXPONENTIAL DISTRIBUTION

As discussed before Poisson r.v. is defined as number of events within an interval of time.

In exponential dist we deal with the length of time or distance between successive events.

The r.v. X equals the distance or time between successive events of a Poisson process with mean number of events $\lambda > 0$ per unit interval is an exponential r.v. with parameter λ. The p.d.f of X is:

$$f(x) = \lambda e^{-\lambda x} \quad \text{for } 0 \leq x < \infty.$$

exp. dist is the prob. dist. that describes the time between events in a Poisson process, i.e. a process in which events occur continuously and independently at a constant average rate.

$X \sim Exp(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0,\ \lambda > 0 \\ 0, & \text{o.w.} \end{cases}$$

$$F(x) = \int_0^x f(u)\,du = \int_0^x \lambda e^{-\lambda u}\,du = \cancel{\lambda} \cdot \frac{e^{-\lambda u}}{-\cancel{\lambda}} \Big|_0^x$$

$$= -e^{-\lambda x} + e^0 = 1 - e^{-\lambda x}$$

Let's check $f(x) \stackrel{?}{=} \dfrac{\partial F(x)}{\partial x}$

$$= \frac{\partial (1 - e^{-\lambda x})}{\partial x} = \lambda \cdot e^{-\lambda x}$$

Let's check whether $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} \lambda . e^{-\lambda x} dx = \int_{0}^{\infty} \lambda . e^{-\lambda x} dx = -e^{-\lambda x} \Big|_0^{\infty} = 0-(-e^{-0}) = 1$$

Ex Let $X \sim Exp(\lambda=2)$ be the legth of the telephone conversation in minutes. Find $P(1 < X \leq 2)$

$$F_X(x) = 1-e^{-\lambda x}$$

$$\lambda = 2 \Rightarrow F(x) = 1-e^{-2x}$$

$$P(1<X\leq 2) = P(X\leq 2) - P(X\leq 1) = F(2)-F(1)$$

$$= 1-e^{-2.2} - (1-e^{-2.1})$$

$$= e^{-4}+e^{-2} = 0.117 \,//$$

Note: The exp. dist. is often used in reliability studies as the model for the time until failure of a device is mostly dealt with.

H.W Suppose X has an exponential dist. with $\lambda=2$, Determine

a) $P(X\leq 0)$ b) $P(X\geq 2)$

c) $P(X\leq 1)$ d) $P(1<X<2)$

e) Find the value of x such that $P(X<x)=0.05$

Try to find the solution by starting with;

$$\int_0^x 2.e^{-2x}dx = 0.05$$

Ex : Let X show the length of a long distance telephone call in minutes. The pdf of X is given by

$$f(x) = \frac{1}{10} e^{-x/10}, \quad x > 0$$

a) Show $f(x)$ is a pdf

b) Find the prob. of a randomly chosen telephone call to be at most 7 minutes and at least 7 minutes respectively.

a) $\forall x$, $f(x) = \frac{1}{10} e^{-x/10} \geq 0$ and $\int_0^\infty f(x)dx = 1$

$$\int_0^\infty \frac{1}{10} e^{-x/10} dx = \frac{1}{10} \int_0^\infty e^{-x/10} dx = \frac{1}{10} \frac{e^{-x/10}}{-\frac{1}{10}} \Big|_0^\infty$$

$$= -e^{-x/10} \Big|_0^\infty = -(\underbrace{e^{-\infty/10}}_{0} - \underbrace{e^{-0/10}}_{1}) = 1$$

b) $$P(X \leq 7) = \int_0^7 \frac{1}{10} e^{-x/10} dx = -e^{-x/10} \Big|_0^7 = -(e^{-7/10} - e^{-0/10})$$

$$= 1 - e^{-7/10}$$

$$= 0.5034$$

$$P(X \geq 7) = \int_7^\infty \frac{1}{10} e^{-x/10} dx = -(e^{-\infty/10} - e^{-7/10})$$

$$= -(-e^{-7/10}) = 0.4966$$

Ex: Suppose that the amount of time one spends in a bank is exponentially dist. with mean 10 min. $(\lambda = 1/10)$. a) What is the prob. that a customer will spend more than 15 minutes in a bank?

b) What is the prob. that a customer will spend more than 15 min. in a bank given that he is still in the bank after 10 minutes?

a) $P(X>15) = e^{-15.\lambda} \Rightarrow e^{-15.\,1/10} = e^{-3/2} = 0.22$

b) $P(X>15 \mid X>10) = P(X>5) = e^{-1/2} = 0.604$ //

Property: Memoryless

$$P(X>s+t \mid X>t) = P(X>s)$$

We can also solve b by using the long way;

$$P(X>15 \mid X>10) + P(X\leq 15 \mid X>10) = 1$$

we can write it because the conditions are the same

$$P(X>15 \mid X>10) = 1 - P(X\leq 15 \mid X>10)$$

$$= 1 - \frac{P(10<X\leq 15)}{P(X>10)}$$

$$= 1 - \frac{F(15)-F(10)}{1-F(10)}$$

$$= 1 - \frac{\cancel{1}-e^{-3/2}-\cancel{1}+e^{-1}}{1-(1-e^{-1})} \Rightarrow 1 - \frac{e^{-1}-e^{-3/2}}{e^{-1}}$$

$$= \frac{\cancel{e^{-1}}-\cancel{e^{-1}}+e^{-3/2}}{e^{-1}} = e^{-1/2}$$ //

Gamma Distribution

The r.v. X with prob. density function

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} \quad \text{for } x > 0$$

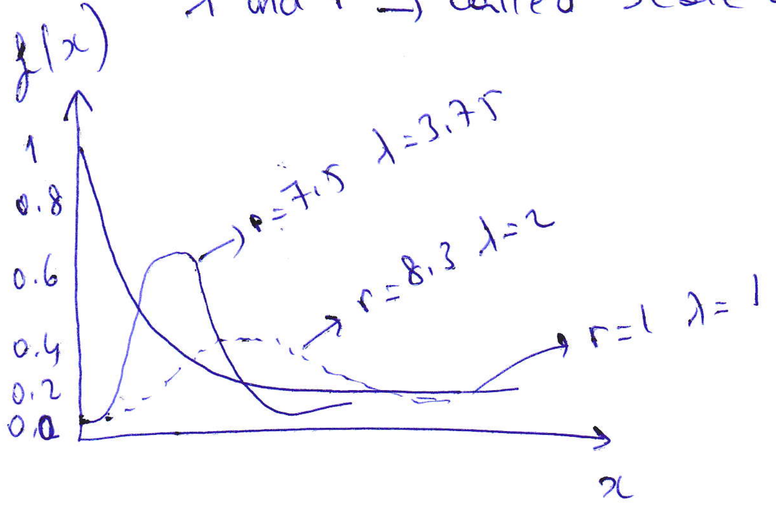
is a gamma distributed r.v. with parameters $\lambda > 0$ and $r > 0$.

Sometimes we find the parameter as; $\lambda = 1/\beta$ and $r = \alpha$

Then;

$$f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, \quad x > 0$$

λ and $r \rightarrow$ called scale and shape parameters, respectively.



$\Gamma(r) \rightarrow$ called the gamma function.

$$\Gamma(r) = \int_0^{\infty} x^{r-1} e^{-x}\, dx \quad \text{for } r > 0.$$

$$\Gamma(r) = (r-1)\,\Gamma(r-1)$$

$$\Gamma(r) = (r-1)!$$

$$\Gamma(1) = 0! = 1$$

$$\Gamma(1/2) = \pi^{1/2}$$

When $r=1$ Gamma converted to exponential distribution.

$$f(x) = \frac{\lambda^1 \cdot x^0 e^{-\lambda x}}{\underbrace{\Gamma(1)}_{1}} = \lambda \cdot e^{-\lambda x}, \quad x > 0$$

Ex: In a company, it is assumed that the electric consumption (thousand kw/hour) is distributed with a gamma distribution with parameters $\alpha = 2$, $\beta = 3$. The circuit station's capacity per a day is 10 thousand kw/hour. Then, what is the prob. of the demand of the company exceeding the capacity of the circuit station in any one of the days?

$\alpha = 2 \quad \beta = 3$

$$f(x) = \begin{cases} \dfrac{e^{-x/3}\, x}{3^2 \cdot \Gamma(2)}, & x > 0 \\ 0, & \text{o.w.} \end{cases}$$

$$P(X > 10) = 1 - \int_0^{10} \frac{x e^{-x/3}}{9} = 1 - \frac{1}{9} \int_0^{10} x \cdot e^{-x/3}\, dx$$

$x = u \qquad e^{-x/3} dx = dv$

$dx = du \qquad \dfrac{e^{-x/3}}{-1/3} = v$

$\Big\}\ u.v - \int v . du$

$$= 1 - \frac{1}{9}\left[x.(-3.e^{-x/3}) \Big|_0^{10} - \int_0^{10} -3.e^{-x/3} dx \right]$$

$$= 1 - \frac{1}{9}\left[10.(-3.e^{-10/3}) - 0 - \left(9 e^{-x/3} \Big|_0^{10}\right) \right]$$

$$= 1 - \frac{1}{9}\left(-1.0702 - 9(e^{-10/3} - e^{-0}) \right]$$

$$= 1 - \frac{1}{9}\left[-1.0702 - 0.3211 + 9 \right] = 0.1546 //$$

NORMAL APPROXIMATION TO BINOMIAL AND POISSON

For large values of n, it is difficult to calculate the probabilities by using the binomial distribution.

The normal approximation is most effective in these cases.

When $n\uparrow$ and p is small.

If X is a binomial r.v. with parameters n and p,

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

is approxiamettly a standard normal r.v. To approxiamate a binomial prob. with a normal distribution, a continuity correction is applied as follows;

$$P(X \leq x) = P(X \leq x + 0.5) \cong P\left(Z \leq \frac{x + 0.5 - n.p}{\sqrt{np(1-p)}}\right)$$

and

$$P(x \leq X) = P(x - 0.5 \leq X) \cong P\left(\frac{x - 0.5 - n.p}{\sqrt{np(1-p)}} \leq Z\right)$$

The approximation is good for $np > 5$ and $n(1-p) > 5$, Otherwise continuity correction can be applied.

Ex: The prob. of the students in a high school to be successful in the university exam is 0.3.

In the high school 100 students got the exam.

a) What is the prob. of 30 students being successful in university exam?

b) Find the prob. of at least 27 of them to pass the exam.

c) Find the prob. of at least 33 and at most 37 of them to pass the exam.

Solution

a) $n=100$ $p=0.3 \Rightarrow n.p = 100 \times 0.3 = 30$

$n.p.q = 100 \times 0.3 \times 0.7 = 21$

$$P(X=x) \cong P\left(\frac{x-0.5-30}{\sqrt{21}} \leq z \leq \frac{x+0.5-30}{\sqrt{21}}\right)$$

$$P(X=30) \cong P\left(\frac{30-0.5-30}{\sqrt{21}} \leq z \leq \frac{30+0.5-30}{\sqrt{21}}\right)$$

$$= P(-0.1091 \leq z \leq 0.1091)$$

$$= 0.5438 - 0.4562 = 0.0876$$

b) $$P(X \geq 27) \cong P\left(z \geq \frac{27-0.5-30}{\sqrt{21}}\right)$$

$$\cong P(z \geq -0.7642)$$

you can also solve it by;

$$P(X \geq 27) = 1 - P(X < 27)$$

$$\cong 1 - P\left(z \leq \frac{(27+0.5)-30}{\sqrt{21}}\right)$$

c) $P(33 \leq X \leq 37)$

$$\cong P\left(\frac{(33-0.5)-30}{\sqrt{21}} \leq Z \leq \frac{(37+0.5)-30}{\sqrt{21}}\right)$$

$\cong P(0.5455 \leq Z \leq 1.6366)$

$\cong P(Z \leq 1.6366) - P(Z \leq 0.5455)$

$\cong 0.9495 - 0.7088$

$\cong 0.2407$ //

Poisson dist. was developed as the limit of a binomial distribution as the number of trials increased to infinity. Consequently, it should not be surprising to find that the normal dist. can also be used to approximate probabilities of a Poisson r.v.

If X is a Poisson r.v. with $E(X) = \lambda$ and $V(X) = \lambda$, then

$$Z = \frac{X - \lambda}{\sqrt{\lambda}}$$

is approximately a standard normal r.v. The same continuity correction used for the binomial dist. can also be applied.

The approxi. is good for $\lambda > 5$.