26.12.18 Normal Distribution

Any distribution with pdf

$$f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & -\infty < x < \infty \\ \end{cases}$$

is called Normal Distribution (where $\sigma$ & $\mu$ are some constants)

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Let $z = \frac{x-\mu}{\sigma}$

$\Rightarrow dz = \frac{1}{\sigma} dx$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2)} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz$$

$= \frac{1}{\sqrt{2\pi}} \times \sqrt{2\pi}$ (since that value equivalent to by result,

$= 1$

$u = \frac{z^2}{2}$ Mean:

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

$= \mu$ (as used in $f(x)$)

$Var = \sigma^2$

Eg. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}$ $\quad \sigma = 1 \quad Var = 1$

$\mu = 1$

To find mgf:

$$E(e^t) = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)\,dx = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx - \frac{(x-\mu)^2}{2\sigma^2}} dx$$

$x = \sigma z + \mu$

Assume $z = \frac{x-\mu}{\sigma}$ $\quad dz = \frac{1}{\sigma} dx$

$$\frac{1}{\cancel{\sigma}\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\sigma z+\mu)} \cdot e^{-\frac{z^2}{2}} \times \cancel{\sigma}\, dz$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t\mu} \cdot e^{t\sigma z} \cdot e^{-z^2/2} dz$$

$a^2 - 2ab$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\sigma t z - z^2/2} dz = \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2\sigma t z)} dz$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 + 2\sigma t z + \sigma^2 t^2 - \sigma^2 t^2)} dz$$

$$= \frac{e^{\mu t + \frac{\sigma^2 t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z - \sigma t)^2} dz$$

$$= \frac{e^{\mu t + \frac{\sigma^2 t^2}{2}}}{\sqrt{2\pi}} \times \sqrt{2\pi} = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$ . mgf

$$M'(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}} \cdot \left(\mu + \frac{\sigma^2 \cancel{2} t}{\cancel{2}}\right)$$

$$M'(0) = e^0 \cdot (\mu + 0)$$
$$= \mu$$

$$M''(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}} (\sigma^2) + (\mu + \sigma^2 t) \cdot e^{\mu t + \frac{\sigma^2 t^2}{2}} \cdot (\mu + \sigma^2 t)$$

At t = 0,

$$M''(t) = e^0 (\sigma^2) + (\mu)(\mu)$$
$$= \sigma^2 + \mu^2$$

$$Var = M''(0) - (M'(0))^2$$
$$= \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

Standard Normal Distribution:

Let $X$ be normal with mean $\mu$, var $\sigma^2$

$$Z = \frac{X - \mu}{\sigma}$$

new random variable = standard normal distribution

$Z$ is called standard normal distribution with mean 0 & variance 1.

Computing mean of $Z$:

Expectation of constant = constant

$$E(Z) = E\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sigma}\left[E(x) - E(\mu)\right]$$

$$= \frac{1}{\sigma}(\mu - \mu) = 0 \text{ for this transformation.}$$

Computing variance of $Z$ = $Var\left(\frac{x-\mu}{\sigma}\right)$

$= \frac{1}{\sigma^2} \, Var(x-\mu) = \frac{1}{\sigma}\left(E((x-\mu)^2) - (E(x-\mu))^2\right)$

$= \frac{1}{\sigma^2}\left(E(x^2) + E(\mu^2) - E(2\mu x) - \overset{0}{\cancel{E(x-\mu)}}\right)$

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

$= \frac{1}{\sigma^2}(\sigma^2) = 1$ for this transformation

Mgf $= \cancel{E}\left(e^{\mu t + \sigma^2 t^2/2}\right) = e^{0 + 1 \cdot t^2/2} = e^{t^2/2}$

$m'(t) = e^{t^2/2} \cdot \frac{\cancel{2}t}{\cancel{2}}$ at $0 \Rightarrow 0$

$m''(t) = e^{t^2/2} + t \cdot e^{t^2/2} \Rightarrow$ At $t=0$, $1+0$

$\sigma^2 = 1 - 0^2 = 1$

Ref Table

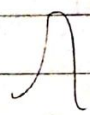
| Distribution (x) | Mean | Var | Mgf | Parameter |
|---|---|---|---|---|
| Uniform disc | $\frac{n+1}{2}$ | $\frac{n^2-1}{12}$ | $\frac{1}{n}\left(\frac{e^t(1-e^{nt})}{1-e^t}\right)$ | $n$ |
| Uniform continuous | $a+b/2$ | $\frac{(a-b)^2}{12}$ | $\frac{1}{t(b-a)}(e^{bt}-e^{at})$ | $a, b$ |
| Binomial | $np$ | $npq$ | $(q+pe^t)^n$ | $(n, p)$ |
| Poisson | $m$ | $m$ | $e^{m(e^t-1)}$ | $m$ |
| Exponential | $1/\lambda$ | $1/\lambda^2$ | $\lambda/\lambda-t$ for $\lambda > t$ | $\lambda$ |
| Normal | $\mu$ | $\sigma^2$ | $e^{\mu t + \frac{\sigma^2 t^2}{2}}$ | $\mu, \sigma, t$ |
| Standard Normal | $0$ | $1$ | $e^{t^2/2}$ | . |

| | Pmf / Pdf | |
|---|---|---|
| UD | $\frac{1}{n}$ $x = 1 \ldots n$, 0 otherwise | $\int ad$ |
| UC | $\frac{1}{b-a}$ $a \le x \le b$, 0 otherwise | |
| B | $p(x) = nC_x p^x q^{n-x}$ $x = 0, 1, 2 \ldots n$ | |
| P | $p(x) = \frac{e^{-m} \cdot m^x}{x!}$ $x = 0, 1, 2 \ldots n$ | |
| E | $\lambda e^{-\lambda x}$ $x \ge 0, \lambda > 0$, 0 for $x < 0$ | |
| N | $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ | $-\infty < x < \infty$ |
| SN | $z = \frac{x - \mu}{\sigma}$ | |

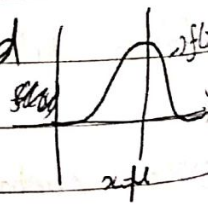
$\sigma^2$

28/12/18 **Properties of normal distribution**

Normal distribution: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$



1) Curve of $f(x)$ is bell-shaped



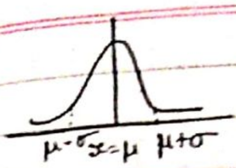

2) $f(x)$ is symmetric about the $x = \mu$ line

3) Maximum probability $= \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(0)} = \frac{1}{\sigma\sqrt{2\pi}}$

4) Limiting case of Binomial is normal, provided $p, q$ are not small.

5) $P(a < x < b) = \int_a^b f(x)\,dx$

Substitute above $f(x)$




0.8413

when $x = \mu$, $z = 0$

$P(\mu-\sigma < x < \mu+\sigma)$

$= P(-\sigma < x-\mu < +\sigma)$

$= P\left(-1 < \frac{x-\mu}{\sigma} < 1\right)$

$= P(-1 < Z < 1)$

$\int_{-1}^{1} \phi(z)\,dz$ where $\phi(z)$ is the pdf of standard normal distribution

$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ ~~dz = 1~~

Now, $= \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-\frac{z^2}{2}}\,dz = 0.9534$

From ~~table~~, table you get 1.6826 $= 2\int_0^1 \phi(x)\,dx$

$= 2 \times 0.3413$

$= \cancel{0.9534}$ $\underline{0.682}$

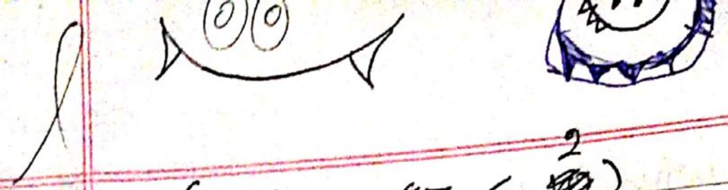
0.4767 $\times$ 2 = 0.9534

- If X is normal variable with mean 30 & s.d 5, find the prob that.
  i) $26 \le X \le 40$
  ii) $X \le 45$
  iii) $|X-30| > 5$

i) $\frac{26-\mu}{\sigma} \le Z = \frac{X-\mu}{\sigma} \le \frac{40-\mu}{5}$

$\downarrow$ $-4/5$

$\downarrow$ $\frac{10}{5} = 2$

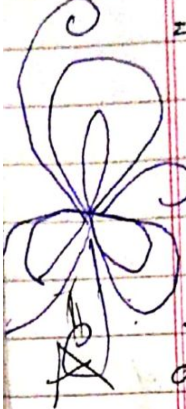


$$P\left(-\frac{4}{5} \le Z \le 2\right)$$

$$= P\left(-\frac{4}{5} \le Z \le 0\right) + P(0 \le Z \le 2)$$

$$= P(0 \le Z \le 4/5) + P(0 \le Z \le 2)$$

$=$



a = 1    b = 10    m = 3    n = 100    30

n/2 = 50    10

n/2 = 25    10

n/2 = 12    10

12/2 = 6    10    7    37

2/1/19

1) Let $X \sim N(\mu, \sigma^2)$    $\mu = 30, \sigma = 5$

Find:

i) $P(26 \le X \le 40)$

ii) $P(X \le 45)$

iii) $P(|X - 30| > 5)$

$30 - 10 = 20$

$20 \le X \le 40$



<u>Answers:</u>

i) $P(26 \le X \le 40)$

$$= P\left(\frac{26-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{40-\mu}{\sigma}\right)$$

$$= P\left(\frac{26-30}{5} \le Z \le \frac{40-30}{5}\right)$$

$$= P\left(-\frac{4}{5} \le Z \le 2\right) = P\left(-\frac{4}{5} \le Z \le 0\right) + P(0 \le Z \le 2)$$

symmetry     Standard normal table

$= P\left(0 \le Z \le \overset{0.8}{\frac{4}{5}}\right) + P(0 \le Z \le 2)$

$= 0.2881 + 0.4772$

$= 0.7653$

iii) $P(|X-30| > 5) = 1 - P(|X-30| \le 5)$

$= 1 - P(-5 \le X-30 \le 5)$

$= 1 - P\left(-1 \le \frac{X-30}{5} \le 1\right)$

$= 1 - P(-1 \le Z \le 1) = 1 - 2P(0 \le Z \le 1)$

$= 1 - 2(0.3413)$

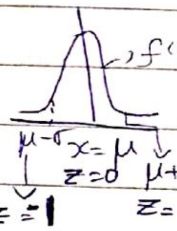
$= 1 - 0.6826 = 0.3174$

Note: By property ⟶

For values of X from

$\mu - \sigma$ to $\mu$ to $\mu + \sigma$,

$Z =$ $-1$ to $0$ to $1$




$\to f(x) = \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

$\frac{\mu+\sigma-\mu}{\sigma} = 1$

ii) $P(\mu - 2\sigma \le X \le \mu + 2\sigma)$

$= P(-2 \le Z \le 2)$

$= 2P(0 \le Z \le 2) = 2 \times 0.4772 = 0.9544$

Beyond ±4 not given in table as 99% of graph is covered within.

Note.

2) Find $P(-\infty < Z \le 1)$

$= P(-\infty \le Z \le 0) + P(0 \le Z \le 1)$

$= \underbrace{0.5}_{\text{covers half of graph}} + 0.3413 = 0.8413$

3) $P(2 \leq z < 2.5)$

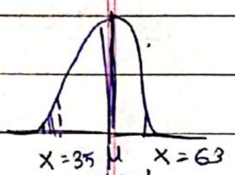
$= P(0 \leq z \leq 2.5) - P(0 \leq z \leq 2)$
$= 0.4938 - 0.4772$
$= 0.0166$

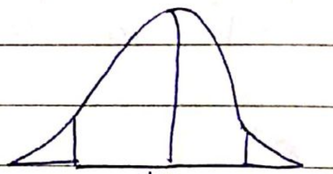
Problem:

In a normal distribution, 7% of the items are under 35, and 89% are under 63. Determine the mean, variance

$P(X \leq 35) = 0.07$
$P(X < \cancel{89}\,63) = 0.89$

$P(X > 63) = 1 - 0.89 = 0.11$




$0.5 - 0.07 = 0.43$ | $0.5 - 0.11 = 0.39$

From table,

$P(\mu \leq x \leq 63) = 0.39$
$P(0 \leq z \leq z_2) =$

$P\left(\frac{\mu-\mu}{\sigma} \leq \frac{x-\mu}{\sigma} \leq \frac{63-\mu}{\sigma}\right)$
$P(0 \leq z \leq z_2) = 0.39$

From table,

$z_2 = \frac{63-\mu}{\sigma} = 1.23$

$1.23\sigma + \mu = 63$

$P(\mu \leq x \leq 35) = 0.43$
$= P\left(0 \leq z < \frac{35-\mu}{\sigma}\right) = 0.43$

From table $-z_1 = \frac{35-\mu}{\sigma} = 1.48$

18/1/19 Normal Distribution Problems

1) If X is a normal distribution with mean 68 & standard deviation 2.5, find

i) $P(66 \le X \le 77)$

ii) $P(69.5 \le X \le 70.5)$

iii) $P(X \ge 72)$

$\mu = 68$

$\sigma = 2.5$

i) $\frac{66-\mu}{\sigma} \le Z \le \frac{77-\mu}{\sigma}$

$\frac{-2}{2.5} \le Z \le \frac{77-68}{2.5} = \frac{9}{2.5}$

$-0.8 \le Z \le 3.6$

0.2881 0.4998 = 7879

$P(-0.8 \le Z \le 0) + P(0 \le Z \le 3.6)$

$= P(0 \le Z \le 0.8) + P(0 \le Z \le 3.6)$

$= 0.7879$

ii) $P(69.5 \le X \le 70.5)$

$\frac{69.5-68}{2.5} \le Z \le \frac{70.5-68}{2.5}$

$= 0.6 \le Z \le 1$

$P(0 \le Z \le 0.6) + P(0 \le Z \le 1)$

$= 0.1156$

iii) $P(X \geq 72)$

$$\frac{72-68}{2.5} = \frac{4}{2.5}$$

1.6

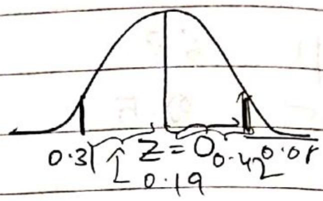
$= 1 - P(X \leq 72)$

$= 1 - 0.9452$

$= 0.0548$

0.9452

2) In a normal distribution, 31 % of the items are under 45 & 8 % are over 64. Find mean & variance of distribution.




$P(X \leq 45) = 0.31$

$P(X \geq 64) = 0.08$

$P(\mu < X < 64) = 0.42$

$$\begin{array}{r} 0.50 \\ 0.08 \\ \hline 0.42 \end{array}$$

$P(0 \leq Z \leq Z_2) = P\left(\frac{\mu-\mu}{\sigma} \leq Z \leq \frac{64-\mu}{\sigma}\right)$

$Z_2 = \frac{64-\mu}{\sigma} = 0.3212$

$$\begin{array}{r} 1.00 \\ 0.08 \\ \hline 0.92 \end{array}$$

$64 = 0.3212\sigma + \mu$

$P(0 \leq Z \leq Z_1) = P(\mu < X < 45)$

$Z_1 = \frac{45-\mu}{\sigma} = 0.1217$

$-Z_1 = \frac{45-\mu}{\sigma} = 0.1217$

$-0.1217\sigma + \mu = 45$

$\sigma = 42.8$

$\mu = 50.2$

19/1/19

# Normal approximation to the Binomial Distribution:

Suppose the no. of successes $X$ ranges from $x_1$ to $x_2$. Then the probability of getting $x_1$ to $x_2$ successes is

$$\sum_{r=x_1}^{x_2} {}^nC_r \, p^r q^{n-r}$$

When $n$ is large, then $p$ is small so Poisson's distribution can be used. (usually)

But for very large $n$ values, both Binomial & Poisson may approach normal distribution curve.

i.e. If $n \to \infty$, using Stirling's formula, Binom. distribution & Poisson, approach(es) Normal Distribu provided $p$ & $q$ do not approach 0 (or cannot be small)

Now,

Pmf for Binom: $P(X=x) = {}^nC_x \, p^x q^{n-x}$

When $p, q \not\to 0$, then we can derive normal distribution here.

Let $\mu = np$ & $\sigma = \sqrt{npq}$

subtract & add some small value

$$P(x_1 \le X \le x_2) = P\left(x_1 - \frac{1}{2} < X < x_2 + \frac{1}{2}\right)$$

done to include $x_1$ & $x_2$

where $X$ = no. of successes

→ To maintain accuracy & include ends, we need to bring in $x_1$ & $x_2$ → less error

because for integral $\int_{x_1}^{x_2}$, $x_2 + 1$ involves next integer, $x_1 - 1$ involves prev integer

we need to include $x_1$ & $x_2$ such that next no. is not involved

Normal is a pdf. Removing endpoints (closed bounds) does not change probability

Converting to Standard normal form,

$$= P\left(\frac{x_1 - \frac{1}{2} - \mu}{\sigma} < Z < \frac{x_2 + \frac{1}{2} - \mu}{\sigma}\right)$$

Substituting for $\mu$ & $\sigma$,

$$= P\left(\frac{x_1 - \frac{1}{2} - np}{\sqrt{npq}} < Z < \frac{x_2 + \frac{1}{2} - np}{\sqrt{npq}}\right)$$

Note: If $p = q = \frac{1}{2}$ & $n$ is small, Binomial $(n, p)$ $\equiv$ Normal $(\mu, \sigma)$.

1) Find the probability that out of 100 ~~q~~ patients between 84 & 95 ~~years of age~~, inclusive, will survive a heart operation, given that chances of survival is 0.9.

$$p = \frac{9}{10} \qquad q = \frac{1}{10}$$

By Binomial formula, if $X$ = no. of survivors

$$P(84 \le X \le 95) = \sum_{r=84}^{95} 100C_r \, (0.9)^r (0.1)^{100-r}$$

Since computation of this is not simple, we use normal distribution.

$$\mu = np = \cancel{100}^{10} \times \frac{9}{\cancel{10}} = 90$$

$$\sigma = \sqrt{npq} = \sqrt{\cancel{100} \times \frac{9}{\cancel{10}} \times \frac{1}{\cancel{10}}} = 3$$

$P(84 \le X \le 95)$ — for more accuracy

can subtract 0.1 — can add 0.1

$= P\left(\frac{84-\frac{1}{2}-90}{3} < Z < \frac{95+\frac{1}{2}-90}{3}\right)$

$= P(-2.16 < Z < 1.833)$

$= P(-2.16 < Z < 0) + P(0 < Z < 1.83)$

$= P(0 < Z < 2.16) + P(0 < Z < 1.83)$

$= 0.4846 + 0.4664$

$= 0.9510$

8 coins are tossed together.

2) Find prob. of getting 1 to 4 heads in a single toss.

$p = \frac{1}{2}$ $q = \frac{1}{2}$

$x_1 = 1$, $x_2 = 4$ $np = 4$

$n = 8$ $npq = 2$ $\sqrt{npq} = \sqrt{2}$

$P(1 \le X \le 4) = P\left(\frac{1-\frac{1}{2}-4}{\sqrt{2}} < Z < \frac{4+\frac{1}{2}-4}{\sqrt{2}}\right)$

$= P(-2.47 < Z < 0.35)$

Ans. $= 0.63$

$\hookrightarrow P(0 < Z < 2.47)$

$+ P(0 < Z < 0.35)$

$= 0.1368 + 0.4932$

$= 0.6300$

3) Find the prob. that by guesswork, a student can correctly answer 25 to 30 qns in an MCQ test consisting of 80 qns.

Each qn has 4 choices & only 1 is correct.

$n = 80$ $x_1 = 25$ $x_2 = 30$

$p = \frac{1}{4}$ $q = \frac{3}{4}$

classmate

$np = \frac{80}{4} = 20$ $npq = \cancel{20}^{5} \times \frac{3}{\cancel{4}} = 15$

$\sqrt{npq} = \sqrt{15}$

$$P\left(\frac{25 - \frac{1}{2} - 20}{\sqrt{15}} < Z < \frac{30 + \frac{1}{2} - 20}{\sqrt{15}}\right)$$

$= P(1.161 < Z < 2.711)$

$= -P(0 < Z < 1.161) + P(0 < Z < 2.711)$

$= \cancel{0.3770} + 0.4966 - 0.3770$

$= \cancel{0.8736}$

$0.9990$, $1.6000$, $0.8736$, $0.127$, $0.1$

$0.4966$, $0.3770$, $0.1196$

should be 0.1196

4) Determine the prob. that getting an even number on face 3 to 5 times in throwing 10 dice together.

$n = 10$ $p = \frac{1}{2}$ $q = \frac{1}{2}$

$x_1 = 3$

$x_2 = 5$

$\mu = np = \frac{10}{2} = 5$ $npq = \frac{20}{4}$

$\sqrt{npq} = \sqrt{\frac{20}{9}}$

$$P\left(\frac{3 - \frac{1}{2} - \cancel{\frac{10}{3}}}{\sqrt{2.5}} < Z < \frac{5 + \frac{1}{2} - \cancel{\frac{10}{3}}}{\sqrt{2.5}}\right)$$

$1.58$ $0.316$

$= P(-\cancel{0.0886} < Z < \cancel{1.2450})$

$= \cancel{0.2123 + 0.4265}$ $0.4429 + 0.1256$

$= \cancel{0.6388}$ $= 0.5685$

Any sample space can have more than 1 r.vars

Date 21/1/19

## Joint Probability Distribution

For a 2-dimensional sample space : eg.

Tossing 3 coins

$S = \{HHH, TTT, HHT, HTH, HTT, THH, THT, TTH\}$

Let random var $X$ = no. of heads in 1st 2 tosses

| $X$ : | 2 | 1 | 0 |
|---|---|---|---|
| $P(X=x)$ : | 1/4 | 1/2 | 1/4 |

Both r.vs are discrete

$Y$ = no. of heads in all 3 tosses

| $Y$ : | 3 | 2 | 1 | 0 |
|---|---|---|---|---|
| $P(Y=y)$ : | 1/8 | 3/8 | 3/8 | 1/8 |

Now, joint pmf is defined as follows:
We take a 2-D random variable

| Y \ X | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 1/8 | 0 | 0 |
| 1 | 1/8 | 2/8 | 0 |
| 2 | ~~1/8~~ 0 | 2/8 | 1/8 |
| 3 | 0 | 0 | 1/8 |

This is the joint pmf of X & Y

$P(X=0, Y=0) = P(X=0 \cap Y=0)$

i.e. $\sum_i \sum_j P(X_i, Y_j) = 1$

The same pmf (joint) can be written as
only non-zero entries

| $P(x,y)$ | $(0,0)$ | $(0,1)$ | ~~$(1,0)$~~ | $(1,1)$ | $(1,2)$ | $(2,$ |
|---|---|---|---|---|---|---|
| $P(X=x, Y=y)$ | $1/8$ | $1/8$ | . | $2/8$ | $2/8$ | $1/8$ |

Joint pmf (joint distribution) is thus defined by the following:

Let X, Y be discrete variables on a sample space S.
Consider function on 2 variables:

$p(x_i, y_j)$. It is said to be joint pmf if:

i) $p(x,y) \geq 0$ for all $x, y$

ii) $\sum_i \sum_j p(x_i, y_j) = 1$

For continuous values of X & Y, we consider the joint pdf such that:

Let X, Y be cont. vars on a sample space S.

Let function on 2 vars $f(x,y)$ be a joint pdf if:

i) $f(x,y) \geq 0$ for all $x, y$

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

To check : eg. $f(x,y) = \frac{9}{4^{x+y}}$ $\quad x = 1, 2, 3 \ldots$
if func is pmf $\quad y = 1, 2, 3 \ldots$

- 1st condition is satisfied
- For 2nd condition,

$$\sum_1^\infty \sum_1^\infty \frac{9}{4^{x+y}} = 9 \sum_1^\infty \sum_1^\infty \frac{1}{4^{x+y}}$$

$$= 9 \sum_{x=1}^\infty \frac{1}{4^x} \sum_{y=1}^\infty \frac{1}{4^y}$$ because y range from 1

$$= 9 \sum_{x=1}^\infty \frac{1}{4^x} \left( \frac{1}{1-\frac{1}{4}} - 1 \right)$$

Now, $1 + x + x^2 \ldots \Rightarrow 4^0$ does not exist

$= \frac{1}{1-x}$ for $|x| < 1$

$$= \sum_{x=1}^\infty \frac{9}{4^x} \left( \frac{4}{3} - 1 \right)$$

$\frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3}$

$\frac{1}{4} \quad \frac{1}{16} \quad \frac{1}{64}$

$\frac{\frac{1}{4}}{\frac{1}{4^2}} \quad \frac{\frac{1}{16}}{\frac{1}{4}} = \frac{1}{4}$

$$= \frac{9}{3} \sum_{x=1}^\infty \left( \frac{1}{4} \right)^x$$

$$= 3 \left( \frac{1}{1 - 1/4} - 1 \right) = 3 \times \left( \frac{4}{3} - 1 \right) = 3 \times \frac{1}{3} = 1$$

Hence it is a pmf

Example: for pdf

$$f(x,y) = \begin{cases} 4xy\, e^{-x^2-y^2} & 0 < x < \infty,\ 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

i) $f(x,y) \geq 0$ always

ii) $$\int_0^\infty \int_0^\infty 4xy\, e^{-x^2-y^2}\, dx\, dy$$

$$= 4 \int_0^\infty y \int_0^\infty x e^{-x^2-y^2}\, dx = 4 \int_0^\infty y e^{-y^2} \left( \int_{-\infty}^\infty x e^{-x^2} dx \right) dy$$

Let $x^2 = u$

$2x\,dx = du$

$$= \int_0^\infty 4y e^{-y^2}\left[\int_0^\infty \frac{1}{2} e^{-u}\,du\right]$$

$$= \int_0^\infty \frac{\cancel{4}^{2}}{\cancel{2}} y e^{-y^2}\left[\frac{e^{-\infty} - e^{0}}{-1}\right]$$

$$= 2 \times \frac{1}{2} = 1$$

- Check whether the following are joint pmf or joint pdf.

i) $f(x,y) = 6x^2y \quad 0<x<1,\ 0<y<1$

$\qquad 0 \quad$ otherwise


ii) $f(x_1,x_2) = x_1 + x_2 \quad 0<x_1<1$

$\qquad 0<x_2<1$

$\qquad 0 \quad$ otherwise

iii) $f(x,y) = \begin{cases} 4xy & 0<x<y \\ & 0<y<1 \\ 0 & \text{otherwise} \end{cases}$

iv) $p(x,y) = \frac{1}{27}(2x+y) \quad x = 0,1,2.$

$\qquad y = 0,1,2.$

22/1/19 Marginal pmf/pdf:

(1) Ex.

| $f(x,y)$ | | X: 0 | 1 | 2 |
|---|---|---|---|---|
| Y | 0 | 1/8 | 0 | 0 |
| | 1 | 1/8 | 2/8 | 0 |
| | 2 | 0 | 2/8 | 1/8 |
| | 3 | 0 | 0 | 1/ |

Let $f_X(x)$ be the marginal pmf.

Adding all y for any x

| | | | | |
|---|---|---|---|---|
| $f_X(x)$ | = $\frac{2}{8}$ | $\frac{4}{8}$ | $\frac{2}{8}$ | marginal distribution on X |

$f_Y(y) = \frac{1}{8} \quad \frac{3}{8} \quad \frac{3}{8} \quad \frac{1}{8}$

This is useful when you need to ~~do~~ perform 2-D conditional probability.

eg. $P(Y=3 \mid X=2) = \frac{P(X \cap Y)}{P(X=2)}$ } not sure check

Definition:

Consider 2 r.vs X, Y with joint pmf/pdf $f(x,y)$. The marginal distribution on ~~Y~~ X is defined by

$f_X(x) = \sum_y f(x,y)$ if X, Y are discrete

If continuous,

$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$

Similarly on Y., $f_Y(y) = \sum_x f(x,y)$ if discrete rVs

$= \int_{-\infty}^{\infty} f(x,y)\,dx$ if continuous rVs

eg. $f(x,y) = \frac{9}{4^{x+y}}$ $\quad x = 1, 2, 3 \ldots$

$y = 1, 2, 3 \ldots$

To find marginal pmf

$f_X(x) = \sum_y f(x,y) = \frac{9}{4^x}\left[\sum_y \frac{1}{4^y}\right]$

Scanned by CamScanner

$\approx$ When $y < 1$, it approaches $\frac{1}{1-x}$

$$\frac{9}{4^x} \times \frac{1}{1-\frac{1}{4}} = \frac{9}{4^x} {}^{3}\times \frac{\cancel{4}}{\cancel{3}}$$

$$= \frac{12}{4^x}$$

Ex. $f(x,y) = 6x^2y \quad 0 < x < 1,\ 0 < y < 1$

$\qquad\qquad\quad 0 \quad$ elsewhere

$6x^2\left[\frac{y^2}{2}\right] \quad (3x^2 = \left[3 \times \frac{x^3}{3}\right] = 1$

This is a valid joint pdf

Marginal on $y = 3x^2$

\#

Note : Let $x, y$ be random variables with a joint pdf or pmf $f(x,y)$

Then $P(a < \cancel{X} < b,\ c < \cancel{X}^{Y} < d)$

$$= \sum_{\cancel{x}=a}^{b} \sum_{y=c}^{d} f(x,y) \text{ if discrete } X, Y$$

$$= \int_a^b \int_c^d f(x,y)\,dy\,dx \text{ is continuous}$$

Cumulative distribution function:

$$F(x,y) = \Pr(X \leq x, Y \leq y)$$

from $-\infty$ onwards possible

$$= \sum_{-\infty}^{x} \sum_{-\infty}^{y} f(x,y) \text{ for discrete } X, Y$$

$$= \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y)\,dy\,dx \text{ for continuous } X, Y$$

Conditional pmf/pdf:

$$P(X=x/Y=y) = \frac{P(X=x, Y=y)}{\underbrace{P(Y=y)}_{\text{marginal pdf for given } Y}}$$ for discrete $X, Y$

$$= \frac{f(x,y)}{f_Y(y)}$$ ~~for discrete~~

When $X, Y$ are continuous,

① Suppose $f(x_1, x_2) = x_1 + x_2$   $0 < x_1 < 1$, $0 < x_2 < 1$

$0$ elsewhere

Find i) $f_{X_1}(x_1)$

ii) $P(X_1 + X_2 \le 1)$

i) $f_{X_1}(x_1) = \int_0^1 x_1 + x_2 \, dx_2 = \left[ x_1 x_2 + \frac{x_2^2}{2} \right]_0^1$

$= x_1 + \frac{1}{2}$

ii) $$\int_0^1 \int_0^{1-x_1} x_1 + x_2 \, dx_2 \, dx_1$$

$$= \int_0^1 x_1 + \left[ \frac{x_2^2}{2} \right]_0^{1-x_1} dx_1 = \int_0^1 x_1 \left[ 1 + x_1^2 - 2x_1 \right]$$

$$\cancel{= \int_0^1 x_1 + x_1^3 - 2x_1^2 \, dx_1} \quad \cancel{x_1\left[1 + x_1^2 - 2x_1\right]}$$

$$\cancel{= \left[ \frac{x_1^2}{2} + \frac{x_1^4}{4} - \frac{2x_1^3}{3} \right]_0^1 = \frac{1}{2\times 6} + \frac{1}{4\times 3} - \frac{2}{3\times 4}}$$

9 − 4

$$\cancel{= \frac{6 + 3 - 4}{12}} = \int_0^1 x_1 - x_1^2 + \frac{1}{2}\left[1 + x_1^2 - 2x_1\right]$$

$$= \frac{x_1^2}{2} - \frac{x_1^3}{3} + \frac{1}{2}\left[x_1 + \frac{x_1^3}{3} - \frac{2x_1^2}{2}\right] = 1/3$$

2) $f(x,y) = 6x^2y \quad 0<x<1,\ 0<y<1$

$\qquad\quad 0 \qquad$ otherwise

$P(0<x<3/4,\ 1/3<y<2)$

$$= \int_0^{3/4}\int_{1/3}^{1} 6x^2y\,dy\,dx$$

$$= \int_0^{3/4} 6x^2\left[\frac{y^2}{2}\right]_{1/3}^{1} = \int_0^{3/4} 6x^2\left[\frac{1\times 9}{2\times 9} - \frac{1}{18}\right]$$

$$= \cancel{6}\times\frac{8}{\cancel{18}3}\int_0^{3/4} x^2$$

$$= \frac{8}{3}\times\left[\frac{x^3}{3}\right]_0^{3/4} = \frac{\cancel{8}}{\cancel{3}}\times\frac{\cancel{27}3}{\cancel{3}\times\cancel{64}8} = \frac{3}{8}$$

3) $f(x,y) = 4xy \quad 0<x<1,\ 0<y<1$

$\qquad\quad 0 \qquad$ otherwise

i) $P\left(0<x<\frac{1}{2},\ \frac{1}{4}<y<\frac{1}{2}\right)$

ii) $P(x<y)$

Answers: ii) $\displaystyle\int_0^1\int_0^y 4xy\,dx\,dy$

$x$ always within $y$ (less than)

& $y$ is always from 0 to

$$= \int_0^1 \cancel{4x}\,\cancel{y^2}\,4y\cdot\left[\frac{x^2}{2}\right]_0^y$$

$$= \int_0^1 4y \times \frac{y^2}{2} = \cancel{2}\left[\frac{y^4}{4/2}\right]_0^1 = \frac{1}{2}$$

iii) $P(X+Y \le 2)$ (or) $P(X \le 2-Y)$

$$\int_0^1 \int_0^{2-y} 4xy\,dx\,dy = \int_0^1 4y \times \left[\frac{x^2}{2}\right]_0^{2-y} =$$

$$= \int_0^1 2y\left[4 + y^2 - 4y\right] = \int_0^1 8y + 2y^3 - 8y^2$$

$$= \left[\frac{8y^2}{2} + \frac{2y^4}{4} - \frac{8y^3}{3}\right]_0^1$$

$$= \frac{8}{2} + \frac{2}{4} - \frac{8}{3} = 4 + \frac{1}{2} - \frac{8}{3} = \frac{24+3-2}{6}$$

$$= \frac{25}{6} \quad 1 + \frac{1}{2} - \frac{4}{3}$$

$$\frac{8\times 6}{2\times 6} \quad \frac{2\times 3}{4\times 3} - \frac{8\times 2}{3\times 2}$$

$$= \frac{3}{2} - \frac{4}{3} = \frac{9-8}{6} = \frac{1}{6}$$

$$\frac{24+6-16}{12}$$

$$= \frac{30-16}{12} = \frac{14}{12}$$

$$\frac{30-16}{14}$$

1

1/1/19 The joint pmf of 2 random variables is given by

$$p(x,y) = \begin{cases} k\,2(x+y) & x = 1,2 \\ & y = 1,2,3 \\ 0 & \text{otherwise} \end{cases}$$

i) Find K

ii) Find marginal pmf of both x & y

iii) Are x, y independent? (Stochastically)

Note: X, Y are independent if $f(x,y) = f_x(x) \cdot f_y(y)$

i) $$\sum_{x=1}^{2} \sum_{y=1}^{3} k \cdot 2(x+y) = 1$$

$$2K \sum_{x=1}^{2} \underbrace{x+1+x+2+x+3}_{3x+6 = 3(x+2)} = 1$$

$$2K \times 3 \sum_{x=1}^{2} (x+2) = 1$$

$$6K(1+2+2+2) = 1$$

$$6K \times 7 = 1$$

$$K = \frac{1}{42}$$

ii) $$P_x(x) = \sum_{y=1}^{3} P(x,y) = \sum_{y=1}^{3} \frac{1}{42} \times 2(x+y)$$

$$= \frac{1}{21} \sum_{y=1}^{3} x+y = \frac{1}{21} \times 3x+6$$

$$= \frac{\cancel{3}(x+2)}{\cancel{21}7} = \frac{x+2}{7}$$

$$P_Y(y) = \sum_{x=1}^{2} P(x,y) = \frac{1}{21}\sum_{x=1}^{2} x+y = \frac{1}{21}(1+y+2+y)$$

$$= \frac{3+2y}{21}$$

iii) $$P_X(x)\cdot P_Y(y) = \frac{x+2}{7} \times \frac{3+2y}{21}$$

$$= \frac{3x+6+2xy+4y}{147}$$

$$\neq P(x,y)$$

Not independent

(2) Joint pdf: $$f(x,y) = \begin{cases} x e^{-x(y+1)} & 0 \le x < \infty,\ 0 \le y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Check if it is a joint pdf

$$\int_0^\infty \int_0^\infty x e^{-x(y+1)}$$

(3) Determine conditional pdf/pmf for the following.

$$P(x,y) = \begin{cases} \frac{1}{30}(2x+y) & x = 1,2 \\ & y = 1,2,3 \\ 0 & \text{otherwise} \end{cases}$$

i) $$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

Correlation coefficients : Applicable for n(data) >

Now, for problem (2)

$$f_Y(y) = \int_0^\infty x e^{-x(y+1)} dx$$

$e^{-x(y+1)}$

$udv = uv - \int v du$ $$= \left[\frac{x e^{-x(y+1)}}{-(y+1)}\right]_0^\infty - \int_0^\infty e^{-x(y+1)} dx$$

$$= -\left[\frac{e^{-x(y+1)}}{y+1}\right]_0^\infty \to 0$$

$$= \frac{1}{y+1} \qquad 0 \le y < \infty$$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{x e^{-x(y+1)}}{\frac{1}{y+1}}$$

$$= x(y+1) e^{-x(y+1)}$$

For problem (3)

$f_Y(y)$ for $\frac{1}{30}(2x+y)$ $\qquad x = 1, 2$
$\qquad y = 1, 2, 3$

$$= \sum_{x=1}^{2} \frac{1}{30}(2x+y)$$

$$= \frac{1}{30}[2+y+4+y] = \frac{1}{30}(6+2y)$$

$$f_{X|Y}(x|y) = \frac{x e^{-x(y+1)} \cdot 15}{y+3} = \frac{y+3}{15}$$

Note:

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$$

(4) The joint cumulative distribution function (CDF) of 2 discrete r.vs X & Y is given by:

$$F(x,y) = \begin{cases} 1/8 & x=1, y=1 \\ 5/8 & x=1, y=2 \\ 1/4 & x=2, y=1 \\ 1 & x=2, y=2 \end{cases}$$

Find $f(x,y)$.

we know that $F(x,y) = Pr(X \leq x, Y \leq y)$

$$f(x,y) = \begin{cases} 1/8 & x=1, y=1 \\ 1/2 & x=1, y=2 \\ 1/8 & x=2, y=1 \\ 2/8 \;\; (\to 1/4) & x=2, y=2 \end{cases}$$

$\frac{1}{4} \times \frac{5}{2 \cdot 8} =$

$P(x=1, y=1) + P(1,2) = 5/8$

$P(X=2, y=1) = P(X=1, Y=1) + P(X=2, y=1) = 1/4$

$\frac{2}{8} + \frac{1}{2} \times \frac{4}{4} = \frac{6}{8}$

(5) $f(x,y) = \frac{1}{2} x^3 y \quad 0 \leq x \leq 2, \; 0 \leq y \leq 1$

$\frac{1}{4 \times 2} \times 2 \; \frac{1}{8}$

i) Are X, Y independent?

$$f_X(X) = \frac{1}{2}\int_0^1 x^3 y \, dy = \frac{x^3}{2} \times \left[\frac{y^2}{2}\right]_0^1 = \frac{x^3}{4}$$

Page

$$f_Y(y) = \frac{1}{2}\int_0^2 x^3 y \, dx = \frac{y}{2} \times \left[\frac{x^4}{4}\right]_0^2$$

$$= \frac{y}{\cancel{2}} \times \frac{\cancel{16}\,\cancel{4}\,2}{\cancel{4}} = 2y$$

Now, $2y \times \frac{x^3}{4} = \frac{1}{2}x^3 y = f(x,y)$

Yes, they are independent.

iii)

$x$

i) $f(x,y) = 10xy^2 \quad 0 < x < y < 1$

$0 \quad$ o therwise

i) Find marginal & conditional pdf

ii) $P(Y > \frac{1}{2} \mid X = 0.25)$

$$f_X(x) = \int_x^1 10xy^2 \, dy$$

$$= 10x \int_x^1 y^2 \, dy = 10x \left[\frac{y^3}{3}\right]_x^1$$

$$= 10x \left[\frac{1}{3} - \frac{x^3}{3}\right] = \frac{10x}{3} - \frac{10x^4}{3}$$

$$= \frac{10}{3} x(1 - x^3)$$

= marginal on X

Conditional mean & variance:

~~E(X|Y) (b|x) \ E(u(X₂)|X₁)~~

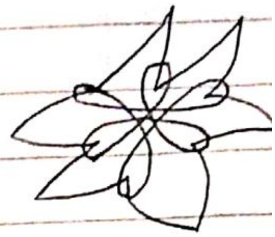
$\mu_{x|y}$

$$E(X|Y) = \int x \, f_{x|y}(x|y)\, dx$$

(or) $$E[u(X_2)|X_1] = \int u(x_2)\, f_{x_2|x_1}(x_2|x_1)\, dx_2$$

$$\sigma^2_{x|y} = Var(X|Y) = E[(X-\mu_{x|y})^2 | Y]$$

$$= \int (x-\mu_{x|y})^2 f_{x|y}\, dx$$



Note:

$$P(a<X<b \mid Y=y) = \int_a^b f_{x|y}(x|y)\, dx$$

Example:

$$f(x,y) = \frac{e^{-x/y}\, e^{-y}}{y} \qquad x>0,\ y>0$$

$$0 \qquad \text{otherwise}$$

Find $E(X|Y=y)$

$\mu_{x|y}$

$$E(X|Y) = \int_{-\infty}^{\infty} x \cdot f_{x|y}(x|y)\, dx$$

Scanned by CamScanner

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{e^{-x/y}e^{-y}}{y \cdot f_Y(y)}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^{\infty} \frac{e^{-x/y}e^{-y}}{y}\,dx$$

$$\frac{e^{-y}}{y}\int_0^{\infty} e^{-x/y}\,dx = \left[\frac{e^{-x/y}}{-1/y}\right]_0^{\infty} \cdot \frac{e^{-y}}{y}$$

$$= \frac{e^{-y}}{\cancel{y}}\left(-(-\cancel{y})\right) = e^{-y} \quad \text{for } y>0$$

Now, $f_{X|Y}(x|y) = \dfrac{e^{-x/y}\cancel{e^{-y}}}{\cancel{e^{-y}}\cdot y}$

$$= \frac{e^{-x/y}}{y} \quad x>0,\ y>0$$

$$E(X|Y) = \int_{-\infty}^{\infty} x \cdot \frac{e^{-x/y}}{y}\,dx$$

M/xi

$$= \frac{1}{y}\int_0^{\infty} x e^{-x/y}\,dx = \frac{1}{y}\left[\frac{xe^{-x/y}}{-1/y} - \frac{e^{-x/y}}{1/y^2}\right]_0^{\infty}$$

$$= \frac{1}{y}\left[-(-y^2)\right] = y \quad \text{for } y>0$$

Scanned by CamScanner

Example 2:

$f(x, y) = 2 \quad 0 < x_1 < x_2 < 1$
$\quad\quad\quad\;\; 0 \quad$ otherwise

Find: 1) $E(X_1 | X_2)$
2) $Var(X_1 | X_2)$
3) $P(0 < X_1 < \frac{1}{2} | X_2 = 3/4)$
4) $P(0 < X_2 < \frac{1}{2})$

Solution: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1\, dx_2$

$$= \int_0^1\int_0^{x_2} 2\, dx_1\, dx_2 = \int_0^1 \left( 2\left[x_1\right]_0^{x_2} \right)$$

$$2\left[\frac{x_2^2}{2}\right]_0^1 = 2 \times \frac{1}{2} = 1 \quad \text{It is a valid joint pdf}$$

$$E(X_1|X_2) = \int_{-\infty}^{\infty} x_1\, f_{x_1|x_2}(x_1|x_2)\, dx_1$$

$$f_{x_1|x_2}(x_1|x_2) = \frac{f(x_1, x_2)}{f_{x_2}(x_2)}$$

$$f_{x_2}(x_2) = \int_0^{x_2} 2\, dx_1 = 2x_2 \quad (0 < x_2 < 1)$$

$$f_{x_1|x_2}(x_1|x_2) = \frac{2}{2x_2} = \frac{1}{x_2}$$

$$E(X_1|X_2) = \int_0^{x_2} x_1 \cdot \frac{1}{x_2}\, dx_1 = \frac{1}{x_2} \times \left[\frac{x_1^2}{2}\right]$$

$$= \frac{x_2}{2}$$

Date
Page

$\mathrm{Var}(X_1 | X_2)$

$= E\left( (X - \mu_{X_1|X_2})^2 \mid X_2 \right)$

$$= \int_0^{x_2} \left(x_1 - \frac{x_2}{2}\right)^2 \cdot f_{x_1|x_2}(x_1|x_2)\, dx_1$$

$$= \int_0^{x_2} \left(x_1^2 + \frac{x_2^2}{4} - x_1 x_2\right) \cdot \frac{1}{x_2}\, dx_1$$

$$\frac{1}{x_2} \int_0^{x_2} x_1^2 + \frac{x_2^2}{4} - x_1 x_2\, dx_1$$

$$\frac{1}{x_2} \left[ \frac{x_1^3}{3} + \frac{x_2^2 x_1}{4} - \frac{x_2 x_1^2}{2} \right]_0^{x_2}$$

$$\frac{1}{x_2} \left[ \frac{x_1^3 \times 4}{3 \times 4} + \frac{x_2^3 \times 3}{4 \times 3} - \frac{x_2^3 \times 6}{2 \times 6} \right]$$

$$\frac{1}{x_2} \left[ \frac{4x_2^3 + 3x_2^3 - 6x_2^3}{12} \right]$$

$$= \frac{x_2^3}{12 \cdot x_2} = \frac{x_2^2}{12}$$

$P(0 < X_1 < 1/2 \mid X_2 = 3/4)$

$$= \int_0^{1/2} f_{X_1|X_2}(x_1|x_2)\, dx_1$$

$$= \int_0^{1/2} \frac{4}{3}\, dx_1 = \frac{4}{3} \times \frac{1}{2} = \frac{2}{3}$$

Scanned by CamScanner

$$p\left(0<x_2<\tfrac{1}{2}\right) = \int_0^{1/2}\int_{0}^{x_2} f(x_1, x_2)\, dx_1\, dx_2$$

$$= \int_0^{1/2}\int_0^{x_2} 2\, dx_1\, dx_2 = \int_0^{1/2} 2\left[x_1\right]_0^{x_2} = \int_0^{1/2} 2(x_2)$$

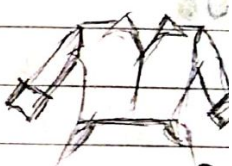
$$= 2\times\left(\frac{x_2^2}{2}\right)_0^{1/2} = 2\times\frac{1}{8} = \frac{1}{4}$$

for $\left(0<x_1<\tfrac{1}{2}\right) = 3/4$

2.2.19 Conditional variance:

$$\sigma_{Y|X} = E\left[(Y-\mu_{Y|X})^2 \mid X\right]$$

$$= E\left[Y^2 \mid X=x\right] - E\left[Y \mid X=x\right]^2$$



Definition

Let $f(x,y)$ be any joint pmf or pdf.

$$\therefore E\left[u(X,Y)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x,y)\cdot f(x,y)\, dx\, dy$$

$$\text{or} \quad \sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} u(x,y) f(x,y)$$

Now,

$$E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\cdot f(x,y)\, dy\, dx$$

$$= \int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f(x,y)\, dy\right) dx = \int_{-\infty}^{\infty} x\, \underbrace{f_X(x)}_{\text{marginal pdf}}\, dx$$

$$E(Y) = \int_{-\infty}^{\infty} y\, f_Y(y)\, dy$$

Let $f_1(x,y)$ be pmf or pdf.

=> Variance $(X) = E(X^2) - (E(X))^2$

↳ marginal

**Eg:** $f(x_1, x_2) = 6x_2 \quad 0 < x_2 < x_1 < 1$

$\qquad\qquad\quad\; 0 \quad$ otherwise

1) Find $E(X_2 | X_1)$
2) Find $E\left(\frac{2X_1}{3}\right)$
3) Find $Var\left(\frac{2X_1}{3}\right)$

$$\int_0^1 \int_{x_2}^1 6x_2\, dx_1\, dx_2 = \int_0^1 \left[6x_2 x_1\right]_{x_2}^1 dx_2$$

$$= \int_0^1 6(x_2 - x_2^2)\, dx_2$$

$$= \left[6\frac{x_2^2}{2} - 6\frac{x_2^3}{3}\right]_0^1$$

$$= \frac{6}{2} - \frac{6}{3} = 3 - 2 = 1$$

$f(x_1, x_2)$ is a joint pd[illegible]

1) To find $E(X_2|X_1)$

$$f_{X_2|X_1}(x_2|x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)}$$

$$f_{X_1}(x_1) = \int_0^{x_1} f(x_1, x_2)\,dx_2 = \int_0^{x_1} 6x_2\,dx_2 =$$

$$\left[\frac{6x_2^2}{2}\right]_0^{x_1} = \frac{6x_1^2}{2} = 3x_1^2 \quad 0 < x_2 < x_1 < 1$$

$$f_{X_2|X_1}(x_2|x_1) = \frac{6x_2}{3x_1^2} \quad 0 < x_2 < x_1 < 1$$

$$= \frac{2x_2}{x_1^2}$$

Now, $E(X_2|X_1) = \int_0^{x_1} x_2 f_{X_2|X_1}(x_2|x_1)\,dx_2$

$$= \int_0^{x_1} x_2 \cdot \frac{2x_2}{x_1^2}\,dx_2 = \frac{1}{x_1^2}\int_0^{x_1} 2x_2^2\,dx_2$$

$$= \frac{1}{x_1^2} \times \left[\frac{2x_2^3}{3}\right]_0^{x_1} = \frac{2\cancel{x_1^3}x_1}{3\cancel{x_1^2}} = \frac{2}{3}x_1$$

$0 < x_2 < x_1 < 1$

2) $E\left(\frac{2X_1}{3}\right)$

Note:
$E(E(X|Y))$
$= E(X)$

$= E(E(X_2|X_1))$

$= E\left(\frac{2X_1}{3}\right) = \frac{2}{3}E(X_1)$

Let $Y$ be $\frac{2}{3}X_1$

Now, $E(Y)$ is to be found

$$E(Y) = E\left(\frac{2}{3}x_1\right) = \int_0^{1}\int_{x_2}^{1} \frac{2}{3} x_1 \times f(x_1, x_2)\, dx_1\, dx_2$$

$$= \int_0^1 \int_{x_2}^1 \frac{2}{\cancel{3}} x_1 \times \cancel{6}^{\,2} x_2\, dx_1\, dx_2$$

$$= 4 \int_0^1 x_2 \times \left[\frac{x_1^2}{2}\right]_{x_2}^1 dx_2 = 4\int_0^1 x_2\left[\frac{1}{2} - \frac{x_2^2}{2}\right] dx_2$$

$$= 4 \int_0^1 \frac{x_2}{2} - \frac{x_2^3}{2}\, dx_2 = 2\left[\frac{x_2^2}{2} - \frac{x_2^4}{4}\right]_0^1$$

$$= 2\left[\frac{1\times 2}{2\times 2} - \frac{1}{4}\right] = 2 \times \frac{1}{4} = \frac{1}{2}$$

Method 2

$G(y) = \Pr(Y \le y)$

3) $\mathrm{Var}\left(\frac{2x_1}{3}\right) = \mathrm{Var}(Y)$

$$= E(Y^2) - (E(Y))^2$$

$$E(Y^2) = \int_0^1 \int_{x_2}^1 \frac{4}{9} x_1^2 \times 6x_2\, dx_1\, dx_2$$

$$= \int_0^1 \left[\frac{4 \times 6x_2}{9} \times \left[\frac{x_1^3}{3}\right]_{x_2}^1\right] = \int_0^1 \frac{24}{9} x_2\left[\frac{1}{3} - \frac{x_2^3}{3}\right] dx_2$$

$$= \frac{24}{9}\left[\frac{x_2^2}{2\times 3} - \frac{x_2^5}{5\times 3}\right]_0^1 = \frac{24}{9}\left[\frac{1}{6} - \frac{1}{15}\right]$$

$= \frac{\cancel{24}^{4}}{\cancel{9}} \left[ \frac{\cancel{9}^{1}}{15 \times \cancel{6}} \right] = \frac{4}{15}$ should be $\frac{1}{60}$

$\frac{4}{15} - \frac{1}{4} = \frac{16-15}{60} = \frac{1}{60}$

$E(X_1) = \iint x_1 f(x_1, x_2)\, dx_1 dx_2$

$= \int_0^1 \int_{x_2}^1 x_1 \cdot 6x_2\, dx_1 dx_2 = \int_0^1 6x_2 \cdot \left(\frac{x_1^2}{2}\right)_{x_2}^1 dx_2$

$= \cancel{6 \int_0^1 x_2 \times \left(\frac{1}{2} - \frac{x_2^2}{2}\right) dx_2} = \cancel{6 \int_0^1 \frac{x_2}{2} - \frac{x_2^3}{2}}$

$= \cancel{6 \left[\frac{x_2^2}{4} - \frac{x_2^4}{8}\right]_0^1} = \cancel{6 \left[\frac{1 \times 2}{4 \times 2} \frac{1}{8}\right]} = \cancel{\frac{1}{8} \times 6} = \cancel{\frac{3}{4}}$

$= \int_0^1 \cancel{6}^{3} x_2 \left(\frac{x_1^2}{\cancel{2}}\right)_{x_2}^1 dx_2 = \int_0^1 3x_2(1-x_2^2)\, dx_2$

$= 3\left(\frac{x_2^2}{2} - \frac{x_2^4}{4}\right)_0^1 = 3\left(\frac{1}{2 \times 2} - \frac{1}{4}\right) = \frac{3}{4}$ check

Compute $E(X_2)$

HW: $f(x_1, x_2) = 8x_1x_2 \quad 0 < x_1 < x_2 < 1$

$\quad 0 \quad$ otherwise

Compute $E(X_1 X_2^2) \Rightarrow 8/21$

$E(X_2) \quad 4/5$

$E(7X_1X_2^2 + 5X_2) \quad \frac{20}{3}$

$$f(x_1, x_2) = \frac{x_1 + 2x_2}{18}$$

when ⟹

$(x_1, x_2) = (1,1), (1,2), (2,1), (2,2)$

~~t~~ 20

0 otherwise

Compute $E(3X_1 - 2X_2)$

Checking,

$$\sum_1^2 \sum_1^2 \frac{x_1 + 2x_2}{18} = \left(\frac{1+2}{18}\right) + \left(\frac{1+4}{18}\right) + \left(\frac{2+2}{18}\right) + \left(\frac{2+4}{18}\right)$$

Expecto patronum!

$$= \frac{3}{18} + \frac{5}{18} + \frac{4}{18} + \frac{6}{18} = \frac{18}{18} = 1$$

It is a valid pmf.

-/2/19 Correlation Coefficient:

Let X, Y be two random variables with mean $\mu_X = E(X)$, $\mu_Y = E(Y)$ & variance $\sigma_X^2 = Var(X)$

$\sigma_Y^2 = Var(Y)$ respectively.

We define covariance of X & Y by

$$\sigma_{XY} = Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Let $f(x, y)$ be joint pdf/pmf,

$$\Rightarrow \sigma_{XY} = E(XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y)$$

Scanned by CamScanner

Let $\rho$, or $\rho_{xy}$ represent the correlation constant.

Now, applying linear property of E,

$$\sigma_{xy} = E(xy) - \mu_x E(y) - \mu_y E(x) + E(\mu_x$$
$$= E(xy) - \mu_x \mu_y - \mu_y \cancel{\mu_x} + \mu_x \cancel{\mu_y}$$
$$= E(XY) - \mu_x \mu_y$$

If x & y are independent,

$$E(xy) = E(x) . E(y)$$
$$= \mu_x \mu_y$$

So $\sigma_{xy} = 0$

But the converse is not true.

If covariance $= 0$, then x & y are not correlated.

Now, $$\rho_{xy} = \frac{cov(x,y)}{\sqrt{Var(x)}\sqrt{Var\,y}} = \frac{\sigma_{xy}}{\sigma_x^2 \sigma_y^2}$$

By this definition,

$$-1 \leq \rho_{xy} \leq 1$$ is the claim.

Now,

$$Var\left(\frac{x}{\cancel{\sigma_x}} + \frac{y}{\sigma_y}\right)$$

Variance is always $\geq$

because

$E(xy) \geq E(x) E($

$$\Rightarrow 0 \leq Var\left(\frac{x}{\sigma_x} + \frac{y}{\sigma_y}\right)$$

$$= E\left(\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y^2}{\sigma_x^2}\right) + \frac{2xy}{\sigma_x \sigma_y}\right)$$

$$-\left(E\left(\frac{x}{\sigma_x} + \frac{y}{\sigma_y}\right)\right)^2$$

$$= \left[\frac{1}{\sigma_x^2} E(x^2) + \frac{1}{\sigma_y^2} E(y^2) + \frac{2E(xy)}{\sigma_x \sigma_y}\right] - \left[\frac{1}{\sigma_x^2} E(x)^2 + \frac{1}{\sigma_y^2} E(y)^2 + \frac{2E(x)E(y)}{\sigma_x \sigma_y}\right]$$

$$= \frac{1}{\sigma_x^2}\left(E(x^2) - E(x)^2\right) + \frac{1}{\sigma_y^2}\left(E(y^2) - E(y)^2\right) + \frac{2}{\sigma_x \sigma_y}\left(E(xy) - E(x)E(y)\right)$$

$$= \frac{Var(X)}{\sigma_x^2} + \frac{Var(Y)}{\sigma_y^2} + \frac{2}{\sigma_x \sigma_y}\left(E(xy) - E(x)E(y)\right)$$

$$= 1 + 1 + 2\rho = 2(1+\rho)$$

$$\Rightarrow 1 + \rho \geq 0 \qquad \rho \geq -1$$

Similarly,

$$0 \leq Var\left(\frac{X}{\sigma_x} - \frac{Y}{\sigma_y}\right)$$

implies $\rho \leq 1$

Therefore, $-1 \leq \rho \leq 1$

Cauchy-Schwartz Inequality in $\mathbb{R}^n$

Take two n-tuples

$$|\vec{a} \cdot \vec{b}| = |a_1 b_1 + a_2 b_2 + \dots a_n b_n|$$
$$\leq (a_1^2 + a_2^2 \dots + a_n^2)^{1/2} (b_1^2 + b_2^2 \dots b_n^2)^{1/2}$$
$$= ||\vec{a}|| \, ||\vec{b}||$$

Example:

Let $f(x,y) = \begin{cases} 25e^{-5y} & 0 < x < 0.2,\ y > 0 \\ 0 & \text{elsewhere} \end{cases}$

Find $\rho_{XY}$

Verification of pdf: $\int_0^{\infty}\int_0^{0.2} 25e^{-5y}\,dx\,dy$

$= \int_0^{\infty}\left(\cancel{25}^{5} e^{-5y} \times \frac{\cancel{2}}{\cancel{10}_{5}}\right) dy = \left[5\,\frac{e^{-5y}}{-5}\right]_0^{\infty}$

$= 0 - (-1) = 1$

$E(XY) = \int_0^{0.2}\int_0^{\infty} xy\,f(x,y)\,dy\,dx = \int_0^{0.2}\int_0^{\infty} xy\,25e^{-5y}\,dy\,dx$

$E(X) = \int_0^{0.2}\int_0^{\infty} x\,f(x,y)\,dy\,dx$

$E(Y) = \int_0^{0.2}\int_0^{\infty} y\,f(x,y)\,dy\,dx$

We can find marginal distribution

$f_X(x) = \int_0^{\infty} f(x,y)\,dy = \int_0^{\infty} 25e^{-5y}\,dy$

$= \cancel{25}^{5} \times \left(\frac{e^{-5y}}{\cancel{-5}}\right)_0^{\infty} = -5(-1) = \begin{cases} 5 & 0 < x < 0.2 \\ 0 & \text{otherwise} \end{cases}$

$$E(x) = \int_0^{0.2} x f_x(x)\, dx$$

Since the distribution is uniform,

$$E(x) = \frac{a+b}{2} = \frac{0.2}{2} = 0.1$$

$$f_Y(y) = \int_0^{0.2} f(x,y)\, dx = 25e^{-5y}(0.2)$$

$$= 5e^{-5y} \quad y > 0$$

$$0 \quad \text{otherwise}$$

$\frac{0.2}{\overline{\text{T}10}}$ $\frac{}{10}$

This is an exponential distribution.

$$E(Y) = \frac{1}{\lambda} = \frac{1}{5} = 0.2$$

$$E(XY) = \int_0^{0.2}\int_0^{\infty} xy \cdot 25e^{-5y}\, dy\, dx$$

$$= 25 \int_0^{0.2} x \left[ \frac{ye^{-5y}}{-5} - \frac{e^{-5y}}{-5^2} \right]_0^{\infty} dx$$

$$= 25 \int_0^{0.2} x \times \frac{1}{5^2}\, dx = \left(\frac{x^2}{2}\right)_0^{0.2} = 0.02$$

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$$

$$= 0.02 - (0.1 \times 0.2)$$

$$= \frac{2}{100} - \frac{1}{10} \times \frac{2}{10} = 0$$

$$\therefore \rho = 0$$

HW: $f(x,y) = \begin{cases} x+y & 0<x<1,\ 0<y<1 \\ 0 & \text{otherwise} \end{cases}$

Find $\rho$ — Ans = $-1/11$

Properties

$Cov(aX, bY) = ab\, Cov(X,Y)$

$Cov(X+a, Y+b) = Cov(X,Y)$ where a, b are constants

If $\rho > 0$, (when $X \uparrow$, Y also $\uparrow$)

$\rho < 0$, (when $X \uparrow$, $Y \downarrow$)

$\rho = 0$ then no relation, (X, Y are uncorrelated)

For, $Y = AX + b$

$Cov(X,Y) = Cov(X, AX+b)$

$= A(Cov(X,X))$

when $A > 0$, then ~~[illegible]~~ $Y = AX+B$ has +ve slope ($Y\uparrow$)

If $A < 0$, slope is -ve. ($Y\downarrow$)

Denominator for $\rho$ is always non-negative

Find $\rho$

$(x,y) = (1,1), (1,2), (1,3), (2,1), (2,2), (2,3)$

| $(x,y)$ | 1,1 | 1,2 | 1,3 | 2,1 | 2,2 | 2,3 |
|---|---|---|---|---|---|---|
| $Pr(X,Y) = (x,y)$ | 2/15 | 4/15 | 3/15 | 1/15 | 1/15 | 4/15 |

$Cov(X,Y) = \dfrac{E(XY) - E(X)E(Y)}{\sqrt{var(X)}\sqrt{var(Y)}}$ all for marginal distribution

$$E(XY) = \sum_{x=1}^{2}\sum_{y=1}^{3} xy\, p(x,y)$$

$$E(X) = \sum_{x=1}^{2}\sum_{y=1}^{3} x\, p(x,y)$$

$$E(Y) = \sum_{x=1}^{2}\sum_{y=1}^{3} y\, p(x,y)$$

$$E(X^2) = \sum_{x=1}^{2}\sum_{y=1}^{3} x^2\, p(x,y)$$

$$E(Y^2) = \sum_{x=1}^{2}\sum_{y=1}^{3} y^2\, p(x,y)$$

Now, $E(XY) = \left(1\times1\times\frac{2}{15}\right) + \left(1\times2\times\frac{4}{15}\right) + \left(1\times3\times\frac{3}{15}\right)$

$+ \left(2\times1\times\frac{1}{15}\right) + \frac{4}{15} + \frac{24}{15} \Big]$

$$\frac{2}{15} + \frac{8}{15} + \frac{9}{15} + \frac{2}{15} + \frac{4}{15} + \frac{24}{15} = \frac{49}{15}$$

| $p(x,y)$ | 1 | 2 | $P_Y(y)$ |
|---|---|---|---|
| 1 | 2/15 | 1/15 | 3/15 |
| 2 | 4/15 | 1/15 | 5/15 |
| 3 | 3/15 | 4/15 | 7/15 |
| $P_X(x)$ | 9/15 | 6/15 | |

$$E(XY) = \frac{2}{15} + 2\times\frac{4}{15} + 3\times\frac{3}{15} + \frac{2}{15} \dots\dots$$

$$\begin{array}{r} 3.2 \\ 15\,\overline{)\,49} \\ \underline{45} \\ 40 \\ \underline{30} \\ 10 \end{array} \qquad = 3.2$$

$$E(X) = \sum_{x=1}^{2} x\, p_X(x)$$

$$= 1\times\frac{9}{15} + 2\times\frac{6}{15} = \frac{9}{15} + \frac{12}{15} = \frac{21}{15}$$

$$E(Y) = \sum_{y=1}^{3} y\, P_Y(y) = 1\times\frac{3}{15} + 2\times\frac{5}{15} + 3\times\frac{7}{15}$$

$$= \frac{3+10+21}{15} = \frac{34}{15}$$

$$E(x^2) = \Sigma x^2 p_x(x)$$
$$= 1\times\frac{9}{15} + 4\times\frac{6}{15}$$
$$= \frac{9}{15} + \frac{24}{15} = \frac{33}{15}$$

$$E(y^2) = \Sigma y^2 p_y(y)$$
$$= 1\times\frac{3}{15} + 4\times\frac{5}{15} + 9\times\frac{7}{15}$$
$$= \frac{3}{15} + \frac{20}{15} + \frac{63}{15} = \frac{86}{15}$$

$$\rho = \frac{cov(xy)}{\sqrt{var(x)}\sqrt{var(y)}}$$

$$= \frac{\frac{49}{15} - \frac{21}{15}\times\frac{34}{15}}{\sqrt{\frac{33}{15} - \left(\frac{21}{15}\right)^2}\sqrt{\frac{86}{15} - \left(\frac{34}{15}\right)^2}}$$

0.24 0.595

$= \frac{2467}{}$ $= 0.\cancel{0126}$ 0.0923

0.377

3.266

HW: A fair die is tossed twice.
Let $X$ = no. of 1s
$Y$ = no. of 3s. Find $\rho_{xy}$

Find $\rho$:

$$f(x,y) = \begin{cases} \frac{1}{3} & \text{if } (x,y) = (0,0), (1,1), (2,2) \\ 0 & \text{elsewhere} \end{cases}$$

$$f(x,y) = \begin{cases} 1/3 & \text{if } (x,y) = (0,2), (1,1), (2,0) \\ 0 & \text{elsewhere} \end{cases}$$

$$f(x,y) = \begin{cases} 1/3 & \text{if } (x,y) = (0,0), (1,1), (2,0) \\ 0 & \text{elsewhere} \end{cases}$$

Inequalities:

1) Cauchy-Schwartz Inequality

$$|E(XY)| \le E(X)E(Y)$$

2) Markov's Inequality

$$Pr(E[u(X) \ge c]) \le \frac{E[u(X)]}{c}$$

for every $c > 0$
provided $E(u(x)) \ge 0$

In particular $P[E(X) \ge c] \le \frac{E(X)}{c}$

3) Chebyshev Inequality

$$\boxed{P(|X-\mu_X| \ge a) \le \frac{\sigma_X^2}{a^2}}$$

Let X be a r.v with finite mean $\mu_X$ & var $\sigma_X^2$

Then this inequality is true for every $a>0$

If $a = k\sigma_X$

$$\boxed{P(|X-\mu| \ge k\sigma_X) \le \frac{1}{k^2}}$$

Upper bounds

$\frac{\sigma_X^2}{a^2}$ or $\frac{1}{k^2}$ is called the upperbounds

Example Let X be a r.v with $\mu = 4$ & $V(X)=9$
Find $P(|X-4| \ge 3)$

$$P(|X-4| \geq 3) \leq \frac{2}{9}$$

When upper bound = 1 then the inequality doesn't hold. i.e. when K = 1.

eg. $P(|X-4| \geq 1) \leq \frac{2}{1} = 2$
↳ not applicable / helpful

$P(|X-1| \geq 3) \leq \frac{2}{9}$ is not necessarily true. because the inequality hasn't been applied. i.e. $E(X) \neq 1$ so it is not applicable

$\frac{1}{3} + \frac{1}{3} + \frac{1}{3} =$ Now, $|x| < a \iff -a < x < a$

$|x| > a \iff x \in (-\infty, -a)$ or $x \in (a, \infty)$

Chebyshev Inequality contd.

Let X be a r.v with finite mean $\mu_x$ & var $\sigma_x^2$,

$$P(|X-\mu| \geq K\sigma) \leq \frac{1}{K^2} \Rightarrow \text{(upper bound)}$$

$$1 - P(|X-\mu| < K\sigma) \leq \frac{1}{K^2}$$

$$P(|X-\mu| < K\sigma) \geq 1 - \frac{1}{K^2} \rightarrow \text{(lower bound)}$$

Proof

To prove $P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}$

We will show that

$$\sigma^2 \geq a^2 P(|X-\mu| \geq a)$$

Assuming that X is continuous, by formula,

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

Now, we know that

$$|x-\mu| \leq a$$

$$\Rightarrow -a \leq x-\mu \leq a$$

$$\mu-a \leq x \leq \mu+a$$

$$\Rightarrow -(a-\mu) \leq x \leq a+\mu$$

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

$$+ \int_{\mu+a}^{\infty} (x-\mu)^2 f(x)\,dx$$

$$= \int_{|x-\mu| \geq a} (x-\mu)^2 f(x)\,dx + \int_{|x-\mu| \leq a} (x-\mu)^2 f(x)\,dx$$

(since $|x-\mu| \leq a$, $-a+\mu \leq x \leq a+\mu$

$$\geq \int_{|x-\mu| \geq a} (x-\mu)^2 f(x)\,dx$$

(since the 2nd term $\int_{|x-\mu| \leq a} (x-\mu)^2 f(x)\,dx$ is always non-negative)

$$\geq a^2 \int_{|x-\mu| \geq a} f(x)\,dx$$

$$\geq a^2 P(|X-\mu| \geq a)$$

We say that the 2nd term is non-negative as follows:

If $f(x) \geq 0$, then integrating over a finite interval $\int_a^b f(x)\,dx$, you get a non-negative value.

i.e. $\mu - a$ to $\mu + a$ is assumed as a finite interval

Example:

$$f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$\mu = \frac{1}{\lambda} = 0.5 \quad (\frac{1}{2})$ Variance $= \frac{1}{\lambda^2} = \frac{1}{4} = 0.25$

From Chebyshev's inequality,

$$P\left(\left|X - \frac{1}{2}\right| \geq \underbrace{\frac{1}{100}}_{a}\right) \leq \frac{1/4}{1/100^2}$$

This number > 1 so it doesn't give the least upper bound

If $a = 100$ $\frac{1}{4} \times \frac{1}{25}$

$$P(|X - 1/2| \geq 100) \leq \frac{1/4}{100^2} = \frac{1}{40000}$$

which is less than 1

So probability can be at most this.

(2) $f(x) = \begin{cases} 1/3 & 1 < x < 4 \\ 0 & \text{otherwise} \end{cases}$

$\mu = \frac{1+4}{2} = \frac{5}{2} = 2.5$ $\sigma^2 = \frac{(4-1)^2}{12} = \frac{9}{12} = \frac{3}{4}$ $= 0.75$

By Chebyshev's Inequality -

$P(|x - 2.5| \geq 5) = \frac{3/4}{5^2} < 1$

↳ upper bound

$P\left(|x - 2.5| \leq \sqrt{\frac{3}{4}} \times 5\right) \geq 1 - \frac{1}{5^2}$ lower bound

(3) If $X$ is a r.v with $E(X) = 3$ & $E(X^2) = 13$

Use Chebyshev's Inequality to determine $P(-2 < x < 8)$

$\mu = 3$ $E(x^2) - (E(x))^2 = \sigma^2$ $13 - 9 = 4$

$P(|x - \mu|) = P(-2-3 < x - 3 < 8 - 3)$

$= P(-5 < x - 3 < 5)$

$\Rightarrow |x - 3| < 5$

$P(|x-3| < \frac{5}{2} \cdot 2) \geq 1 - \frac{1}{(5/2)^2}$ $1 - \frac{4}{5^2}$

Note: $P(|x - \mu| \geq a) \leq \frac{\sigma^2}{a^2}$

$P(|x - \mu| \leq a) \geq 1 - \frac{\sigma^2}{a^2} = 1 - \frac{4}{25} = \frac{21}{25}$

$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

& $P(|X-\mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$

$\therefore P(|X-\mu| < 5) = P(|X-\mu| \leq \frac{5}{2} \times 2)$

$\geq 1 - \frac{1}{(5/2)^2} = \frac{21}{25}$ which gives the same answer

A symmetric coin is tossed 1600 times. What is the probability that the head will be shown up > 1200 times?

$\mu = np = 1600 \times \frac{1}{2} = 800$

$\sigma^2 = npq = 1600 \times \frac{1}{2} \times \frac{1}{2} = 800 \times \frac{1}{2} = 400$

If X is the no. of heads, we compute

$P(|X-\mu| > 1200-\mu)$

$P(|X-800| > 1200-800)$

$P(|X-800| > 400) \leq \frac{1}{400}$

$\downarrow$

$K \times 20$

Notes (Probability)

Date / /
Page

STUDY BUDDIES

8/2/19 Note: Let $f(x,y)$ be joint pmf or pdf

(i) $E[x+y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\, f(x,y)\,dx\,dy = \underbrace{E(x)+E(y)}_{\text{possibly}}$

(ii) $E([x+y]^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)^2 f(x,y)\,dx\,dy$

(iii) $Var(x+y) = E([x+y]^2) - (E(x+y))^2$

**Proof:**

Central Limit Theorem:

Let $X$ be any distribution with the pdf/pmf $f(x)$, with mean $\mu$ & variance $\sigma^2$
Consider $n$ observations.

Let $X_1, X_2 \ldots X_n$ be the $n$ observations having the above mean $\mu$ & variance $\sigma^2$
(all of the distribution $X_i$ has the same mean & var)

Let $S_n = X_1 + X_2 + \cdots + X_n$

$E(S_n) = E(X_1 + X_2 \ldots X_n)$
$= E(X_1) + E(X_2) + \cdots + E(X_n)$
$= n\mu$

To

$\Rightarrow Var(S_n) = n\sigma^2$

Assume $Z_n = \dfrac{S_n - E(S_n)}{\sqrt{Var(S_n)}} = \boxed{\dfrac{S_n - n\mu}{\sqrt{n\sigma^2}}}$

$$\boxed{Z_n = \frac{S_n - n\mu}{\sqrt{n}\,\sigma}}$$

Dividing num & denom by $n$,

$$\boxed{Z_n = \frac{\frac{S_n}{n} - \mu}{\frac{1}{\sqrt{n}}\sigma}} \Rightarrow \boxed{Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}$$

$\frac{\sqrt{n}}{n}$

Date ___/___/___
Page ______
STUDY BUDDIES

where $\overline{X} = \frac{S_n}{n} = \frac{X_1 + X_2 \dots X_n}{n}$

The central limit theorem says that
$\lim_{n \to \infty} Z_n$ is a normal distribution with mean $\sim N(0,1)$
0 & var 1 if $n \geq 30$

i.e. $Z_n$ is a standard normal distribution, as long as
$\underbrace{n \geq 30}$ (i.e. for sufficiently large n).
Statistically large samples.

Example:

Let X be random variable with below pdf

$f(x) = \frac{1}{3} \quad 1 < x < 4$
$\quad\quad 0 \quad$ otherwise

Let $X_1, X_2 \dots X_n$ be a random variable with pdf $f(x$

$\mu = E(X) = \frac{4+1}{2} = \frac{5}{2} \qquad \sigma^2 = var(x) = \frac{(b-a)^2}{12}$

$= \frac{9}{12} = \frac{3}{4} \qquad \sigma = \frac{\sqrt{3}}{2}$

Find $P(108 \leq S_n \leq 126)$ Let $n = 48$

$\hookrightarrow P\left(\frac{108 - 48(5/2)}{\sqrt{48} \cdot \frac{\sqrt{3}}{2}} \leq \frac{S_n - n\mu}{\sqrt{n}\,\sigma} \leq \frac{126 - 48(5/2)}{\sqrt{48} \cdot \frac{\sqrt{3}}{2}}\right)$

$P\left(\frac{-12 \times 2}{12} \leq Z \leq \frac{6 \times 2}{\cancel{12}\, 6}\right)$

$P(-2 \leq Z \leq 1)$

$= 0.4772 + 0.3413$

$= \underline{0.8185}$

Scanned by CamScanner

8) A normal population has mean of 0.1 & ~~an~~ std dev of 2.1.

Find prob that mean of sample size 900 will be -ve.

Let $\overline{X}$ be the sample mean

$n = 900$

To find $(\overline{X} < 0)$

$$P\left(\frac{\overline{X} - \cancel{n}\mu/\cancel{n}}{\sigma/\sqrt{n}} < \frac{0-0.1}{2.1/\sqrt{900}}\right) \qquad \overline{X} = \frac{S_1 + S_2 \ldots + S_n}{n}$$

$$P\left(Z_n < \frac{-0.1 \times 30}{2.1}\right) = P(Z_n < -1.43)$$

From table & by central limit theorem,

$$P(Z_n < -1.43) = 0.5 - 0.4236$$
$$= 0.0764$$

(3) If the mean breaking strength of Cu wire is 575 lbs with standard dev 8.3, how large a sample must be used in order that there will be 1 chance in 100 that the mean breaking strength of sample is < 572 lbs

$\mu = 575$

$\sigma = 8.3$

Let $\overline{X}$ = sample mean

$$P(\overline{X} < 572) = \frac{1}{100}$$

$$P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < \frac{572 - 575}{8.3/\sqrt{n}}\right) = 0.01$$

$$P\left(Z_n < \frac{-3\sqrt{n}}{8.3}\right) = 0.01$$

$0.05 - x = 0.01$

Date / /
Page

STUDY BUDDIES

From central limit theorem, $Z_n$ follows normal distribution $N(0,1)$

$$-\frac{3\sqrt{n}}{8.3} = 0.004$$

$$1 - P\left(Z > -\frac{3\sqrt{n}}{8.3}\right) = 0.01$$

$$P\left(Z > -\frac{3\sqrt{n}}{8.3}\right) = 0.99$$

$$0.5 + P\left(0 > Z > -\frac{3\sqrt{n}}{8.3}\right) = 0.99$$

$$P\left(0 < Z < +\frac{3\sqrt{n}}{8.3}\right) = 0.49$$

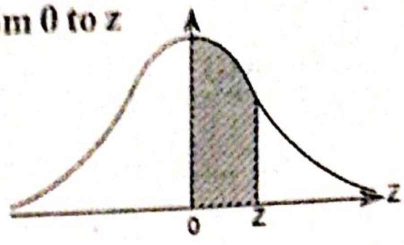
$$+\frac{3\sqrt{n}}{8.3} = 2.33 \qquad n = 42$$

(4) The guaranteed average life of a certain type of electric bulb is 1500 hours w/ $\sigma = 120$ hours. It is decided to sample the output so as to ensure that 95% of bulbs do not fall short of the guaranteed average by more than 2%. What is the min. sample size? Ans: 44

**Table – 3**

**Areas under the Standard Normal Curve from 0 to z**

$$Z = \frac{x - \mu}{\sigma}$$




| z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| 0.1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0754 |
| 0.2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| 0.3 | .1179 | .1217 | .1256 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| 0.4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |
| 0.5 | .1916 | .1950 | .1985 | .2019 | .2054 | .2088 | .2123 | .2157 | .2190 | .2224 |
| 0.6 | .2258 | .2291 | .2324 | .2357 | .2389 | .2422 | .2454 | .2486 | .2518 | .2649 |
| 0.7 | .2580 | .2612 | .2642 | .2673 | .2704 | .2734 | .2764 | .2794 | .2823 | .2852 |
| 0.8 | .2881 | .2910 | .2939 | .2967 | .2996 | .3023 | .3051 | .3078 | .3106 | .3133 |
| 0.9 | .3159 | .3186 | .3212 | .3238 | .3264 | .3289 | .3315 | .3340 | .3365 | .3389 |
| 1.0 | .3413 | .3438 | .3461 | .3485 | .3508 | .3531 | .3554 | .3577 | .3599 | .3621 |
| 1.1 | .3643 | .3665 | .3686 | .3708 | .3729 | .3749 | .3770 | .3790 | .3810 | .3830 |
| 1.2 | .3849 | .3869 | .3888 | .3907 | .3925 | .3944 | .3962 | .3980 | .3997 | .4015 |
| 1.3 | .4032 | .4049 | .4066 | .4082 | .4099 | .4115 | .4131 | .4147 | .4162 | .4177 |
| 1.4 | .4192 | .4207 | .4222 | .4236 | .4251 | .4265 | .4279 | .4292 | .4306 | .4319 |
| 1.5 | .4332 | .4345 | .4357 | .4370 | .4382 | .4394 | .4406 | .4418 | .4429 | .4441 |
| 1.6 | .4452 | .4463 | .4474 | .4484 | .4495 | .4505 | .4515 | .4525 | .4535 | .4545 |
| 1.7 | .4654 | .4564 | .4573 | .4582 | .4591 | .4599 | .4608 | .4616 | .4625 | .4633 |
| 1.8 | .4641 | .4649 | .4656 | .4664 | .4671 | .4678 | .4686 | .4693 | .4699 | .4706 |
| 1.9 | .4713 | .4719 | .4726 | .4732 | .4738 | .4744 | .4750 | .4756 | .4761 | .4767 |
| 2.0 | .4772 | .4778 | .4783 | .4788 | .4793 | .4798 | .4803 | .4808 | .4812 | .4817 |
| 2.1 | .4821 | .4826 | .4830 | .4834 | .4838 | .4842 | .4846 | .4850 | .4854 | .4857 |
| 2.2 | .4861 | .4864 | .4868 | .4871 | .4875 | .4878 | .4881 | .4884 | .4887 | .4890 |
| 2.3 | .4893 | .4896 | .4898 | .4901 | .4904 | .4906 | .4909 | .4911 | .4913 | .4916 |
| 2.4 | .4918 | .4920 | .4922 | .4925 | .4927 | .4929 | .4931 | .4932 | .4934 | .4936 |
| 2.5 | .4938 | .4940 | .4941 | .4943 | .4945 | .4946 | .4948 | .4949 | .4951 | .4952 |
| 2.6 | .4953 | .4955 | 4956 | .4957 | .4959 | .4960 | .4961 | .4962 | .4963 | .4964 |
| 2.7 | .4965 | .4966 | .4967 | .4968 | .4969 | .4970 | .4971 | .4972 | .4973 | .4974 |
| 2.8 | .4974 | .4975 | .4976 | .4977 | .4979 | .4978 | .4979 | .4979 | .4980 | .4981 |
| 2.9 | .4981 | .4982 | .4982 | .4983 | .4984 | .4984 | .4985 | .4985 | .4986 | .4986 |
| 3.0 | .4987 | .4987 | .4987 | .4988 | .4988 | .4989 | .4989 | .4989 | .4990 | .4990 |
| 3.1 | .4990 | .4991 | .4991 | .4991 | .4992 | .4992 | .4992 | .4992 | .4993 | .4993 |
| 3.2 | .4993 | .4993 | .4994 | .4994 | .4994 | .4994 | .4994 | .4995 | .4995 | .4995 |
| 3.3 | .4995 | .4995 | .4995 | .4996 | .4996 | .4996 | .4996 | .4996 | .4996 | .4997 |
| 3.4 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4998 |
| 3.5 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 |
| 3.6 | .4998 | .4998 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.7 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.8 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.9 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 |