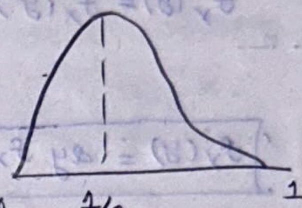
② $f_{\theta|X}(\alpha|0) = \dfrac{6\alpha(1-\alpha)(1-\alpha)}{\int_0^1 6\alpha(1-\alpha)^2\, d\alpha} = \dfrac{6\alpha(1-\alpha)^2}{\int_0^1 6\alpha(1+\alpha^2-2\alpha)\, d\alpha}$

$= 12\alpha(1-\alpha)^2$

'favouring towards the result'



'favouring towards $X = x$'

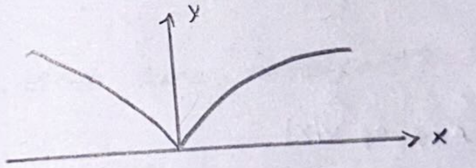
observe 0 → pushes our beliefs that the Bias of the coin is low.

$p(\theta|0)$ will be lower.

3. Derived distribution

Know X, want $Y = g(X)$

1) CDF of Y: $F_Y(y) = P(Y \le y)$
2) Sub $Y \to X$ using $g(\cdot)$
3) ~~dest~~ before rewrite as CDF of X



$X \sim N(0, 1)$ ⟶ standard normal

$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

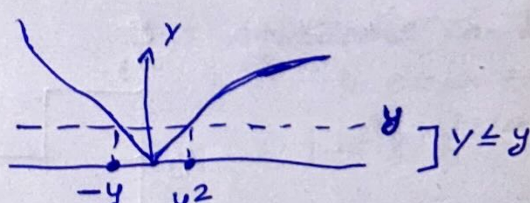
let $Y = g(X)$,

pdf of Y = ?

$$g(t) = \begin{cases} -t, & t \le 0 \\ \sqrt{t}, & t > 0 \end{cases}$$

soln:

$F_Y(y) = P(Y \le y)$



∴ left side: $g(t) = -t$

∴ we have $Y \le y$ | $\sqrt{y^2} = y$

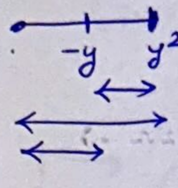
∴ $-(-y) = y$

Idea:

Derive CDF of y in terms of X.

$f_Y(y) = P(Y \le y) = P(-y \le X \le y^2)$

$= P(X \le y^2) - P(X \le -y)$



$\boxed{F_Y(y) = F_X(y^2) - F_X(-y)}$

$f_Y(y) = f_X(y^2) \cdot (2y) - f_X(-y)(-1)$

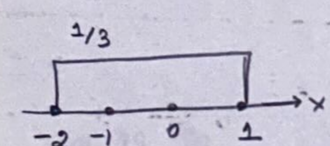
"chain rule"

$\boxed{f_Y(y) = 2y\, f_X(y^2) + f_X(-y)}$

$$f_Y(y) = \begin{cases} 2y\left(\frac{1}{\sqrt{2\pi}} e^{-y^4/2}\right) + \left(\frac{1}{2\pi} e^{-y^2/2}\right) & \to \text{only } y>0 \\ 0 & \text{otherwise} \end{cases}$$

Pdf

a) $f_X(x) = \begin{cases} 1/3, & -2 < x \le 1 \\ 0, & \text{otherwise} \end{cases}$



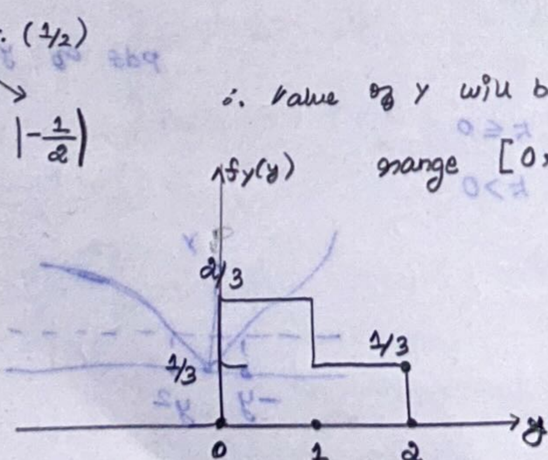
, pdf of $|X|$

$Y = |X|$

Sol: say: (1/2)

$\left|\frac{1}{2}\right|$ $\left|-\frac{1}{2}\right|$

∴ value of Y will be (on) range $[0, 2]$



$P(0 \le Y \le 1) = P(-1 \le X \le 0) + P(0 \le X \le 1)$

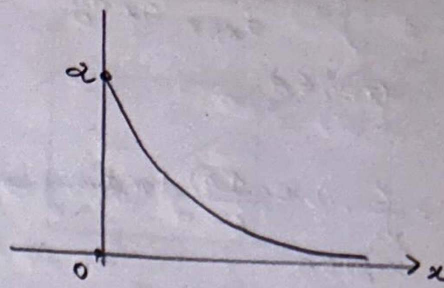
$= 1/3 + 1/2 = 2/3$

$P(1 \le Y \le 2) = P(-2 \le X \le -1)$

$= 1/3$

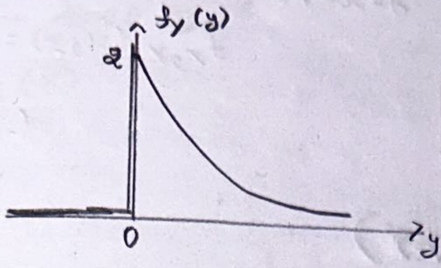
$$f_Y(y) = \begin{cases} 2/3 & 0 \le y \le 1 \\ 1/3 & 1 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$$

b) $f_X(x) = \begin{cases} \dots \\ 0, & o.w \end{cases}$



soln: $Y = |X| = X$

$$f_Y(y) = \begin{cases} 2e^{-2y}, & y > 0 \\ 0, & o.w. \end{cases}$$



x is already +ve

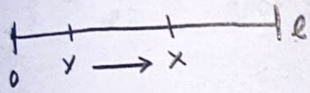
$|X| = X.$

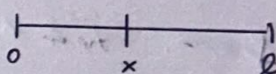
iii) $P(|X| = x) = P(X = x) + P(X = -x)$

$f_Y(y) = f_X(y) + f_X(-y) \rightarrow$ continuous

$\boxed{\begin{matrix} x = +y \\ x = -y \end{matrix} \quad |X| = y}$

<u>Ambulance travel time</u>





How much time takes to respond?

Accident
* Somewhere b/w 0 and l occurs
* Ambulance is b/w 0 & l.

$P(T \leq t)$?

$\therefore T$ is a function of X & Y.

$$T = \frac{|Y - X|}{v} \quad v \rightarrow \text{speed of ambulance.}$$

$$P(T \leq t) = P\left(\frac{|Y - X|}{V} \leq t\right)$$

$$= P\left(-Vt \leq Y - X \leq Vt\right)$$

$\therefore$ Ambulance can be in either side of accident

$$= P\left(X - Vt \leq Y \leq X + Vt\right)$$

$X - Vt \downarrow$ It returns

$X + Vt \downarrow$ It further moves

$$P(T \leq t) = P(X - Vt \leq Y \leq X + Vt)$$

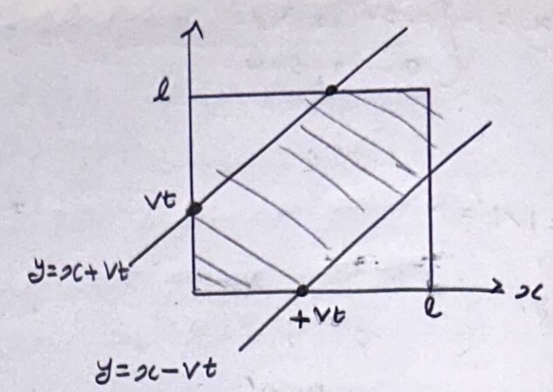
Assume: $A: \{(x, y) : x - vt \leq y \leq x + vt$

$x \in [0, l]$

$y \in [0, l]$

$\downarrow$ set of all x & y.

$$= \int_{(x,y)\in A} f_{X,Y}\, dx\, dy$$



$$= \int_{(x,y)\in A} \frac{1}{l^2}\, dx\, dy$$

$$= \frac{1}{l^2}\int dx\, dy$$

$f_{X,Y}(x,y) = \frac{1}{l^2}$

$(\because \int = 1)$

$$= \frac{1}{l^2}\left[\text{Area of square} - 2\Delta^{les}\right]$$

$$= \frac{1}{l^2}\left(l^2 - 2\times\frac{1}{2}\times(l-vt)\times(l-vt)\right)$$

$$= \frac{2vt}{l} - \frac{(vt)^2}{l^2} \qquad \left[0 \le t < 1/v\right]$$

$$F(t) = P(T\le t) = \begin{cases} 0 & \text{if } t<0 \\ \frac{2vt}{l} - \frac{(vt)^2}{l^2} & \text{if } 0 \le t < l/v \\ 1 & \text{if } t \ge l/v \end{cases}$$

$$f(t) = \begin{cases} \frac{2v}{l} - \frac{2v^2 t}{l^2}, & 0 \le t < \frac{l}{v} \\ \end{cases}$$

$y = x - vt$

As x ranges from 0 to l
y ranges from 0 to l

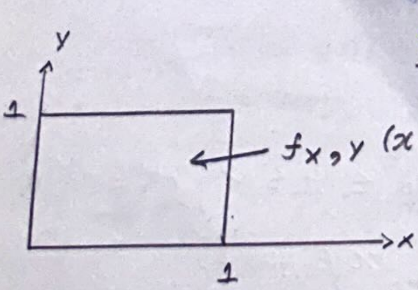
Time: when X & Y are same zero.

maximum: X ——— Y

$\therefore \frac{\text{length}}{\text{velocity}}$

Lecture-11

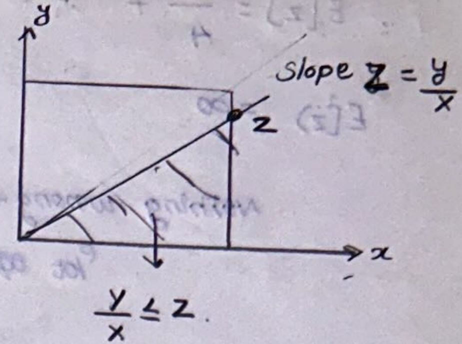
Derived Pdfs, Covariance



$f_{X,Y}(x,y) = 1$ (uniform)

$Z = g(X,Y) = Y/X$

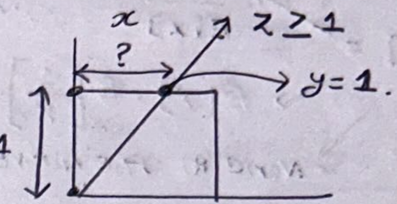
$$F_Z(z) = P_Z\left(\frac{Y}{X} \le z\right)$$



$$F_Z(z) = P_Z\left(\frac{Y}{X} = z\right) = \frac{1}{2} \times z \times 1$$

$$= \frac{z}{2} \le 1$$

(one case)



$F_Z(z)$ = whole Area - Area of Δ^{le}

$= 1 - \frac{1}{2} \times 1 \cancel{x} (x)$

$= \frac{y}{x} = z$

$\frac{1}{x} = z$

$\boxed{\cancel{x} = \frac{1}{z}}$

$z \to$ slope

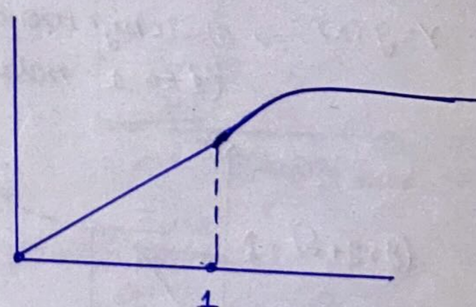
$= 1 - \frac{1}{2} \times \frac{1}{z}$

$= 1 - \frac{1}{2z}$

$$F_Z(z) = \begin{cases} z/2 & z \le 1 \\ 1 - \frac{1}{2z} & z \ge 1 \\ 0 & \text{otherwise} \end{cases}$$

$\therefore$ x & y are +ve random variables

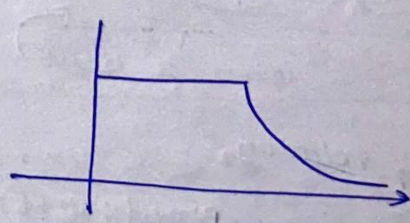
$\therefore$ No way of -ve ratio.



Pdf - density

$$f_Z(z): \begin{cases} 0, & z < 0 \\ 1/2 & z \le 1 \\ -\left(\frac{1}{2}\right)\frac{z^{-2}}{\cancel{2}} \cancel{x - 1} = \frac{1}{2z^2} & z \ge 1. \end{cases}$$

$\downarrow$ Derivative



Expected value

$$E[Z] = \int_0^1 z \cdot \frac{1}{2}\, dz + \int_1^\infty z \cdot \frac{1}{2z^2}\, dz$$

$$= \frac{1}{2}\left(\frac{z^2}{2}\right)_0^1 + \int_1^\infty \frac{1}{z}\, dz.$$

$$\frac{d}{dz}\left(-\frac{1}{2} z^{-1}\right) = -\frac{1}{2}(-1) z^{-2}$$

$$= \frac{1}{2z^2}$$

$E[z] = \frac{1}{4} + \ln(\infty) - \ln(1)$

$E[z] = \infty$

Nothing wrong → ∴ Tail of function falls slow

"lot of R.V has ∞ Expectation"

$Z = \frac{Y}{X}$

$E[z] = \rightarrow \frac{E[Y]}{E[X]}$

$\rightarrow E[Y] \cdot E\left[\frac{1}{Y}\right]$

None of the above

what will be $E[z]$ → average

$\frac{Y}{X}$ → Not linear, so $E[z] \neq \frac{E[Y]}{E[X]}$

∴ Y & X are independent

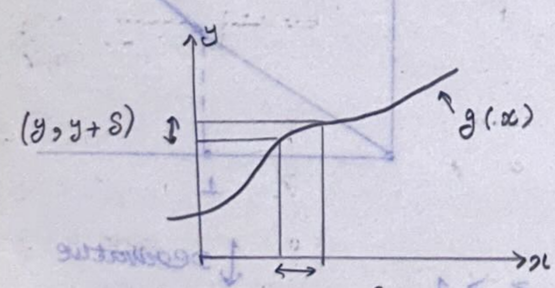
Y & $\frac{1}{X}$ are independent

$E[z] = E[Y] \cdot E[1/X]$

Special case - Don't go through CDF :-

$Y = g(x)$ → strictly monotonic

(1 to 1 relation b/w X & Y)



∴ when x falls, $g(x)$ falls in $(y, y+\delta)$

∴ relating,

$Prob(y, y+\delta) = Prob(x, x+\delta)$

little intervals → density.

P(little interval) = density × δ (length)

$(x, x+\delta)$, $(y, y+\delta)$

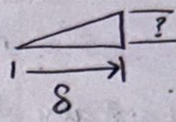
can't necessarily be equal

Interval of y

'Depends on the slope'

* $x \le X \le x + \delta x$
* $g(x) \le Y \le g(x+\delta)$
* APP

$g(x) \le Y \le g(x) + \delta \left| \frac{dg}{dx}(x) \right|$



slope = $\frac{y}{x}$

$\boxed{y = \text{slope} \times \delta}$

$$\boxed{\delta f_X(x) = \delta f_Y(y) \left| \frac{dg}{dx}(x) \right|}$$

$y = g(x)$

Now:

Formula: density of y as a function of y.

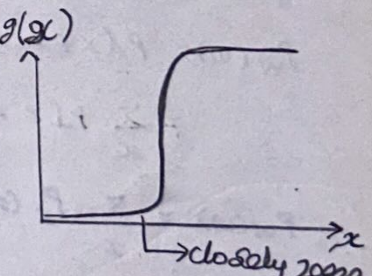
$y = x^3$, $\boxed{g(x) = x^3 = y}$ $\boxed{x = y^{1/3}}$

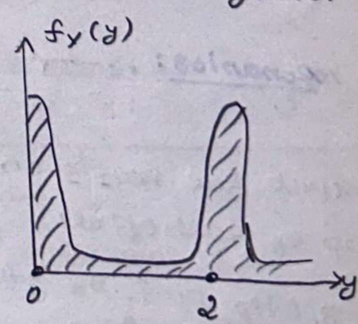
$f_X(x) = f_Y(y) \cdot \text{slope}$

$f_X(x) = f_Y(y) \cdot 3x^2$

$f_Y(y) = f_X(y^{1/3}) \cdot \frac{1}{3}\left(\frac{1}{y^{1/3}}\right)^2$

$f_Y(y) = f_X(y^{1/3}) \cdot \frac{1}{3}\left(\frac{1}{y^{2/3}}\right)$

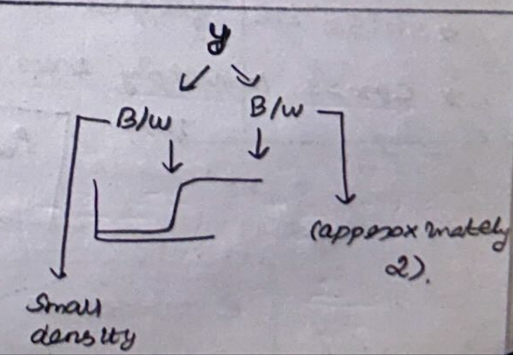




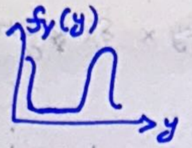
If x is uniform, what will be y?
(reasonable flat density)

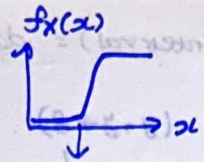
From x → mostly the prob will be zero
(or)
will be near 2.

rarely in intermediate.



Inference: Flat Slope:





Large prob of y

$\therefore$ Flat slope $\rightarrow$ large prob.

(Inv. prop to slope)

$\therefore f_Y(y) = f_X(y^{1/3}) \cdot \frac{1}{3}\left(\frac{1}{y^{1/3}}\right)$ $\rightarrow$ Inverse slope.

This case: works for one to one monotonically $\uparrow$ or $\downarrow$ cases

In monotonically $\downarrow$ case $\rightarrow$ Take absolute values of slope

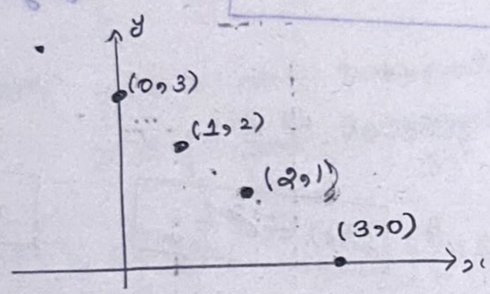
---

The distribution of $X+Y$

$W = X+Y$, X, Y independent.

warmup: look at discrete case

Let $W=3$



$$P_W(w) = P(X+Y=w)$$
$$= \sum_x P(X=x) \cdot P(Y=w-x)$$

$\therefore$ If $x=3$
$y = w-3$
$= 3-3$
$\boxed{y=0}$

$$P_W(w) = \sum_x P_X(x)\, P_Y(w-x)$$

$\downarrow$

Convolution formula

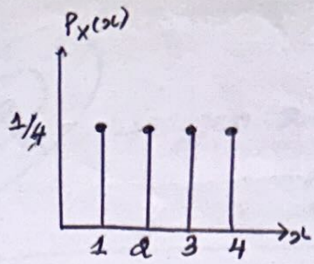
Takes 2 pmfs, produces new pmf

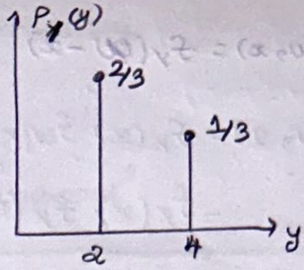
(Transformation)

mechanics:

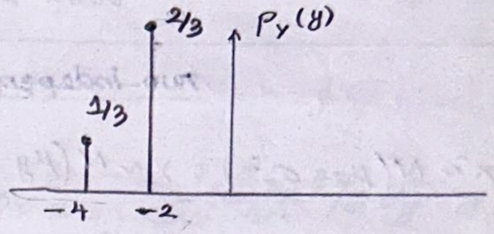
* put the pmf's on top of each other
* Flip pmf of y
* shift the flipped pmf by w
* cross multiply then add.

$$\boxed{P_W(w) = \sum_x P_X(x)\, P_Y(w-x)}$$





(con)

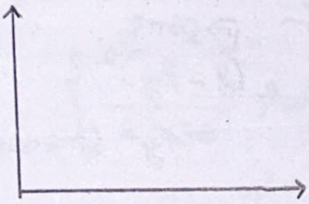


when $W=0$,

$$\sum_x P_X(x)\, P_Y(-x)$$

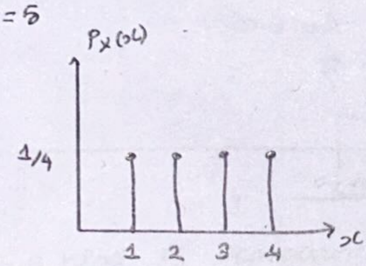
<u>cross multiply</u>:

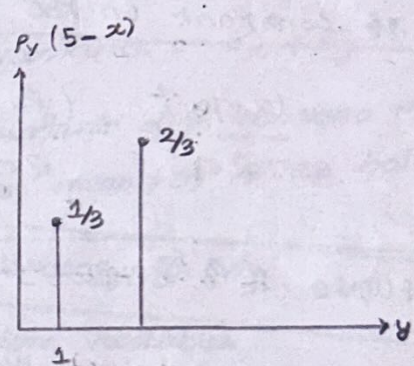
Step 1

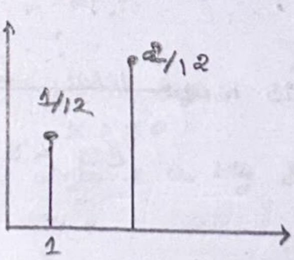


zero no over lap b/w $P_X(x)\ P_Y(-x)$

when $w=5$







Continuous Case

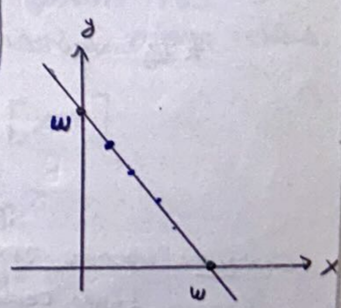
$w = x+y$, x, y independent.

$\boxed{\text{pmf by pdf, sum by } \int \text{gral}}$

★ $f_{W|X}(w|x) = ?$

<u>Actually w is y:</u>

when $x=0$, $w=y$



$x+y=w$

$\boxed{\text{Function } w \text{ is in the height } y \text{ in each } x\text{'s}}$

* $f_{W|X}(w, x) = f_Y(w - x)$

* $f_{W,X}(w, x) = f_X(x)\, f_{W|X}(w|x)$

$= f_X(x)\, f_Y(w - x)$

$$\boxed{f_W(w) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(w - x)\, dx}$$

↳ Just translated version of discrete case.

Two independent normal r.v.s

$X \sim N(\mu_x, \sigma_x^2)$; $Y \sim N(\mu_y, \sigma_y^2)$

$$f_{X,Y}(x, y) = f_X(x)\, f_Y(y)$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{ \frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2} \right\}$$

PDF is constant on the ellipse where:

$$\frac{(x-\mu_x)^2}{2\sigma_x^2} + \frac{(y-\mu_y)^2}{2\sigma_y^2} \text{ is constant}$$

$\boxed{\text{* Ellipse is a circle when } \sigma_x = \sigma_y}$

centered at μ_x & μ_y

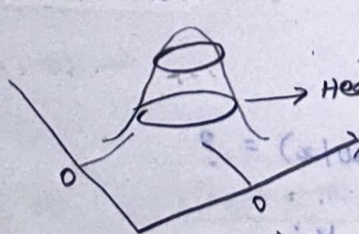
X is more likely to spread out than Y.

"Big X's have small prob that of small Y"

(variance X) > (variance Y)

$\sigma_x > \sigma_y$ (or) $\sigma_y > \sigma_x$

* If variance of Y is bigger, Y takes bigger values.



→ Here variances of X & Y are same.

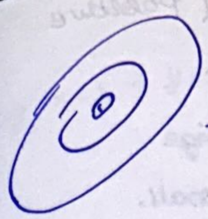
(wide)

→ zero mean

2-d bell,

Contour: where JPDF takes constant.

(outline)



→ Bigger x → Bigger y

"Dependent normal"

Sum of Independent normal r.v.s

- $X \sim N(0, \sigma_x^2)$, $Y \sim N(0, \sigma_y^2)$

$$W = X + Y$$

$$f_W(w) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(w-x)\, dx$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} \int_{-\infty}^{\infty} e^{-x^2/2\sigma_x^2} \cdot e^{-(w-x)^2/2\sigma_y^2}\, dx$$

(Algebra) $= c e^{-\gamma w^2}$

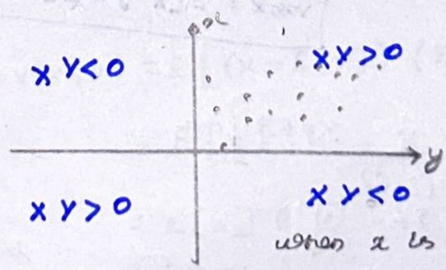
Conclusion: W is normal

* mean $= 0$, variance $= \sigma_x^2 + \sigma_y^2$

* same argument for non zero mean case.
  (when mean $\neq 0$) → This holds.

Covariance (Dependency of RV)

* Some kind of dependence b/w random variables



$E[XY] > 0$ (most of points in +ve quadr)

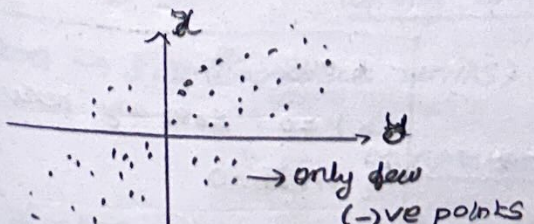
when x is big y is big too.

$$\text{cov}(X, Y) = E\Big[(X - E[X])\,(Y - E[Y])\Big]$$

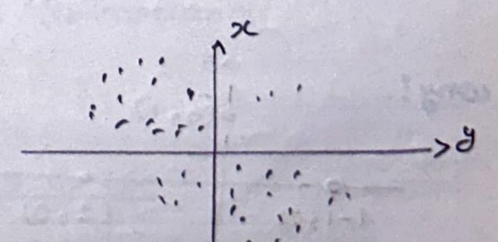
deviation of x from its mean value & y from its mean value.

zero mean case

zero mean case: $\text{Cov}(X, Y) = E[XY]$



$\text{cov}(x, y) > 0$



$\text{Cov}(x, y) < 0$

positive covariance → indicates → systematic relation (positive association b/w two)

+ve covariance → when x is large, y is large

-ve covariance → when x is large, y is small.

$$\mathrm{Cov}(X, X) = E\left[(X - E[X])^2\right]$$

$$= \text{variance}(X)$$

Indendent

$\mathrm{Cov}(x, y) = 0$

$$\mathrm{cov}(X, Y) = E[XY] - E[X]E[Y]$$

Covariance: Useful: Calc. variance of sum of R.V

Independent

$$\mathrm{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{var}(X_i)$$

Dependent

$$\mathrm{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{var}(X_i) + \sum_{(i,j)\, i \neq j} \mathrm{cov}(X_i, X_j)$$

Inference: Zero means (say)

$$\boxed{\mathrm{var}\, X = E[X^2] - (E[X])^2}$$

$$\mathrm{var}(X) = E\left[(X_1 + X_2 + \dots X_n)^2\right]$$

$$= E\left[\sum (X_i^2 + \dots) + \sum_{\substack{(i,j) \\ i \neq j}} X_i X_j\right]$$

Expected value of this is variance

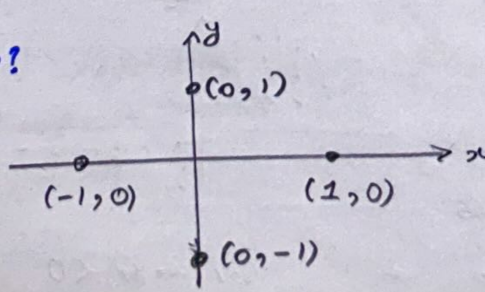
covariance.

Independent: $X_i X_j = 0$

$$E[XY] = E[X]E[Y]$$

covariance = 0 for Independent

why?



4 outcomes

$XY = 0$ for any here.

mean = 0

$\therefore$ covariance$(x, y) = 0$

Proop

Knowledg $X=1 \rightarrow Y=0 \rightarrow$ Dependent

$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])]$$

$$= E[XY] - E[XE[Y]] - E[E[X]Y] + E[E[X]E[Y]]$$

$$= E[XY] - E[X]E[Y] - E[X]E[Y] + E[Y]E[X]$$

$\therefore E[X] \rightarrow$ constant

$$= E[XY] - E[X]E[Y]$$

$cov(aX+b, Y) =$ (zero means)

$$= E[(ax+b)y] = aE[XY] + bE[Y]$$

$$= aE[XY]$$ (zero means) $\downarrow$ Individual.

$$cov(X, Y+Z) = E[X(Y+Z)]$$

$$= E[XY] + E[XZ]$$

$$= Cov(x,y) + cov(x,z)$$

Independent

$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])]$$

$$= E[XY]$$ Zero mean

$$= E[X]E[Y]$$ Independent

$$= 0$$

Ind. case

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

---

wrong units (covariance)

Variance $\rightarrow$ S.D (correct units)

Correlation Co.eff

* Dimensionless version of covariance.

$$\rho = E\left[\frac{(X - E[X])}{\sigma_X} \cdot \frac{(Y - E[Y])}{\sigma_Y}\right]$$

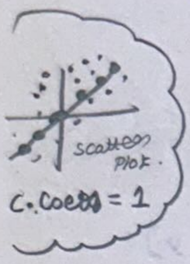
→ standardising
ind. std deviation.

$$= \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

∴ units of both are same.

Use: measure of strength of the correlation co.eff

$[-1, 1] \to$ range



* $-1 \le \rho \le 1$
* $|\rho| = 1 \iff (X - E[X]) = c\,(Y - E[Y])$

(Linearly related) → when it has extreme value.

(Complete Correlation)

Zero correlation: lack of systematic relation.

1 or −1: strong association.

Recitation

$X, Y \to$ Juliet at time (X → Romeo)

$X, Y \sim \exp$ with parameter λ

Interest: Difference b/w two arrivals

$$Z = X - Y \Rightarrow Z = X + (-Y)$$

$$\boxed{\begin{array}{c} W = X + Y \\ f_W(w) = \int_{-\infty}^{\infty} f_X(x) f_Y(w - x)\,dx \end{array}}$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) \cdot f_{-Y}(z - x)\,dx$$

$$\boxed{f_{-Y}(z - x) = f_Y(x - z)}$$

ours case

$z = x - y$

$\boxed{y = z - x}$

$$P(-Y = 1) = P(Y = -1)$$

$$P(-Y = -1) = P(Y = 1)$$

$$f_X(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \ge 0 \end{cases}$$

$$= \int_{-\infty}^{\infty} f_X(x)\, f_Y(x - z)\,dx.$$

$$f_Y(y) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda y}, & x \ge 0. \end{cases}$$

$$= \int_0^{\infty} \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(x - z)}\,dx$$

$$= \int_0^{\infty} \lambda^2 e^{-\lambda x - \lambda x + \lambda z}\,dx$$

$$= \lambda^2 e^{\lambda z} \int_0^\infty e^{-2\lambda x}\, dx$$

$$= \frac{\lambda^2 e^{\lambda x}}{\lambda} \left(-\frac{1}{2} e^{-2\lambda x}\right)_0^\infty$$

$$= \frac{\lambda}{2} e^{\lambda z} \left(\cancel{0} + \frac{1}{2}\right)$$

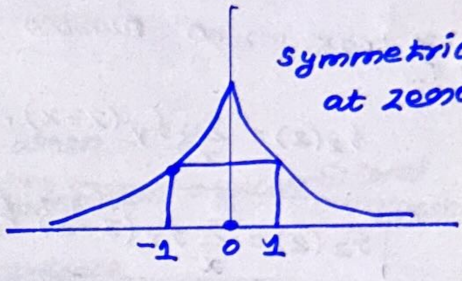
$$= \frac{\lambda}{2} e^{\lambda z}, \quad z < 0$$

$z = x - y$

It may vary from (−)ve to (+)ve

$\boxed{-\infty < z < \infty}$

$$f_Z(z) = \begin{cases} \frac{\lambda}{2} e^{\lambda z}, & z < 0 \\ \frac{\lambda}{2} e^{-\lambda z}, & z \geq 0 \end{cases}$$



symmetric at zero.

$\therefore x - y$ & $y - x$ has same distribution

$z = x - y$

$-z = y - x$

$\therefore x$ & y are ind, identically distributed

$$\boxed{f_Z(z) = f_Z(-z)}$$

$$f_Z(z) = \begin{cases} \frac{\lambda}{2} e^{\lambda z}, & z < 0 \\ \frac{\lambda}{2} e^{-\lambda z}, & z \geq 0. \end{cases}$$

$0 \leq x \leq 1$

$0 \leq x \leq 1$

$-1 \leq z \leq 1$

Sum of discrete & continuous variable

X, Y ind, $X \rightarrow$ discrete r.v, PMF $P_X(x)$

$Y \rightarrow$ continuous r.v, Pdf $f_X(x)$

pdf of $Z = X + Y$, $f_Z(z) = ?$

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

$= P(X + Y \leq z)$ $\rightarrow$ Total prob theorem by conditioning

$= \Sigma P(X + Y \leq z \mid X = x)\, P(X = x)$

$= \Sigma P(X + Y \leq z \mid X = x)\, P_X(x)$

$= \sum_x P(X + Y \leq z \mid X = x)\, P_X(x)$

$\therefore X$ & Y are independent

$= \sum_x P(Y \leq z - x \mid X = x)\, P_X(x)$

(No use - conditioning)

$$= \sum_x P(Y \le Z-x)\, P_X(x)$$

$$= \sum_x F_Y(z-x)\, P_X(x)$$

Derivative:

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \frac{d}{dz} \sum_x F_Y(z-x)\, P_X(x)$$

$$f_Z(z) = \sum_x \frac{d}{dx} F_Y(z-x)\, P_X(x)$$

If x took on ∞ number of values — caution of using.

$$f_Z(z) = \sum_x f_Y(z-x) \cdot (1)\, P_X(x)$$

$$\boxed{f_Z(z) = \sum_x f_Y(z-x)\, P_X(x)}$$ → similar to convolution formula.

Lecture: 12: Iterated expectations; Sum of a Random no. of R.V

Conditional expectation: $E[X|Y]$

$$\boxed{E[X|Y=y] = \sum_x x\, P_{X|Y}(x|y)}$$

Integral & pdf in Continuous case.

* Stick length l → chose Y, chose X.

$E[X|Y=y] = \frac{y}{2}$ (number) (may take any value - based on experiment)

$E[X|Y] = \frac{Y}{2}$ (r.v) → "Before experiment we don't know y' as well as X. (Itself can be assumed as R.V)

when a number is gn, it will be a number.
[Before experiment $\frac{Y}{2}$]

Law of Interrelated Iterated expectation

$$\boxed{\begin{aligned} g(Y) &= E[X|Y] \\ E[X|Y=y] &= g(y) \end{aligned}}$$

$g(y)$ → number (gn)

* $E[E[X|Y]] = \sum_y E[X|Y=y]\, P_Y(y) = E[X]$

$$E[g(Y)] = \sum_y g(Y) \cdot P_Y(y)$$

$$= \sum_y E[X|Y] \cdot P_Y(y)$$ → Total expectation theorem

$$= E[X]$$

$$\boxed{E_X(x) = \sum_y P_Y(y) \cdot E[X|Y]}$$

↓ Summing all y's

In stick example

$$E[X] = E\left[E[X|Y]\right] = E\left[\frac{Y}{2}\right] = \frac{1}{2}E[Y]$$

$$E[X|Y] = \frac{Y}{2}$$

$$E[Y] = \frac{l}{2}$$

$$= \frac{1}{2}\left(\frac{l}{2}\right) = \frac{l}{4}.$$

conditional Expectation (In a conditional world)

$$\text{var}(X|Y=y) = E\left[\left(X - E[X|Y=y]\right)^2 \mid Y=y\right]$$

$$= E\left[\left(X - E[X|Y=y]\right)^2 \mid Y=y\right] \rightarrow \text{R.V (we don't know X is)}$$

R.V
(we don't know - depends on Y)

* $\text{var}(X|Y)$: a r.v
with value $\text{var}(X|Y=y)$ when $Y=y$.

"Expected value of a conditional expectation = unconditional expectation"

Law of total variance

$$\text{var}(X) = E[\text{var}(X|Y)] + \text{var}(E[X|Y])$$

"In the case of var(X), taking expectation of cond. Exp is not true"

Proof

$\rightarrow$ $\text{var}(X) = E[X^2] - (E[X])^2$

$\rightarrow$ $\text{var}(X|Y) = E[X^2|Y] - (E[X|Y])^2$ $\rightarrow$ conditioned on Y

$\rightarrow$ $E[\text{var}(X|Y)] = E[X^2] - E\left[(E[X|Y])^2\right]$ $\longrightarrow$ 3

$\rightarrow$ $\text{var}(E[X|Y]) = E\left[(E[X|Y])^2\right] - E\left[E[X|Y]\right]^2$ $\rightarrow$ 4

$$\boxed{\text{var}(X) = E[X^2] - (E[X])^2}$$

$$= E\left[(E[X|Y])^2\right] - (E[X])^2$$

Sum of R.H.S 3 & 4

$$E[X^2] - E\left[(E[X|Y])^2\right] + E\left[(E[X|Y])^2\right] - (E[X])^2$$

$$\boxed{var(x) = E[x^2] - (E[x])^2}$$

$\therefore \; var(x) = E[var(x|y)] + var(E[x|y]) \longrightarrow$ Proved

---

Example: Two sections : 1, 2   section:1 : 10 students, section 2 = 20 students

x: quiz score, y = sections

$E[x|y=1]: \frac{1}{10}\sum_{i=1}^{10} x_i = 90$, $E[x|y=2]: \frac{1}{20}\sum_{i=11}^{30} x_i = 60$

$E[x] = \frac{1}{30}\sum_{i=1}^{30} x_i = \frac{(90\times10)+(60\times20)}{30} = 70 \longrightarrow$ Expected quiz score if I pick a student

↓ Expectation of quiz scores

1 class, 2 class

i) $E[x|y=1] = 90$, $E[x|y=2] = 60$

$$E[x|y] = \begin{cases} 90, & \text{w.p. } 1/3 \\ 60, & \text{w.p. } 2/3 \end{cases}$$

↓ R.V

$E[E[x|y]] = \left(\frac{1}{3}\cdot 90\right) + \left(\frac{2}{3}\cdot 60\right) = 70 = E[x]$

↓ Two ys

Average class 1, class 2

Expectation (Average C1 + C2)

= Average quiz marks

$var(E[x|y]) = \frac{1}{3}(90-70)^2 + \frac{2}{3}(60-90)^2$

$= \frac{600}{3} = 200 \longrightarrow (1)$

$\therefore$ we don't put $E[x|y]$ without a value (conditioned)

$E[x|y=1] = \sum_x x \cdot P_x(x)$

$E[E[x|y]] = \sum_y \underbrace{E[x|y=y]}_{x} \cdot P_y(y)$ ↳ $\therefore$ In terms of y we writing

---

$var(E[x|y]) \rightarrow$ how each section is varying from total average.

Conditional variances

Inside each sections:

$\frac{1}{10}\sum_{i=1}^{10}(x_i-90)^2 = 10$, $\frac{1}{20}\sum_{i=1}^{20}(x_i-60)^2 = 20$

$var(x|y=1) = 10$, $var(x|y=2) = 20$

$$var(x|y) = \begin{cases} 10, & \text{w.p } 1/3 \rightarrow \text{Section 1} \\ 20, & \text{w.p } 2/3 \rightarrow \text{Section 2.} \end{cases}$$

$$E[var(x|y)] = \frac{1}{3} \cdot 10 + \frac{2}{3} \cdot 20$$

$$= \frac{50}{3}$$

↳ e
"I want the entire unconditioned universe"

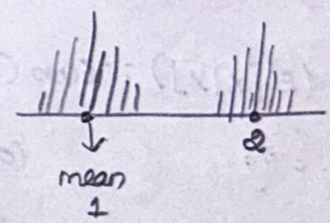
$$var(x) = E[var(x|y)] + var(E[x|y])$$

$var(E[x|y]) \rightarrow$ Random variable
as y value is not provided.
we can talk about expected value.

$$= \frac{50}{3} + var(E[x|y])$$

↳ variances of individual sections.

$$var(E[x|y]) = E[E[x|y]^2] - (E[x])^2$$



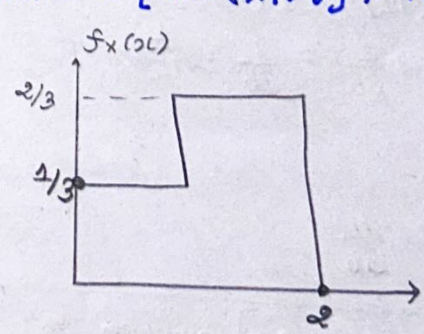
$$= \left(\frac{1}{3}\right)90^2 + \left(\frac{2}{3}\right)60^2 - 4900$$

$$= 200$$

[variability b/w sections]

$$= \frac{50}{3} + 200 = \frac{650}{3} = 416.66$$

$$var(x) = E[var(x|y)] + var(E[x|y])$$

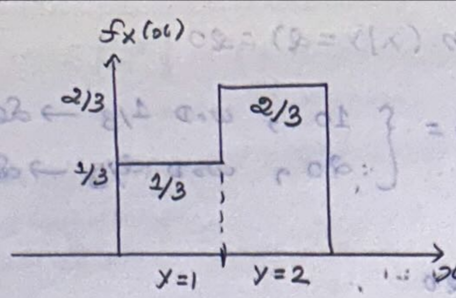


variance?

Solu:

Define : Divide & Conquer.

Trick



(Defining a R.V)

$E[X|Y=1]$ = uniform $= \frac{1}{2}$, $E[X|Y=2] = 3/2$ (b/w 1 & 2)

$var(X|Y=1) = \frac{(b-a)^2}{12} = \frac{1}{12}$, $var(X|Y=2) = 1/12$

i) $E[X] = E\big[E[X|Y]\big]$

$= \left(\frac{1}{3}\right)\cdot\left(\frac{1}{2}\right) + \left(\frac{2}{3}\right)\left(\frac{3}{2}\right)$

$= \frac{1}{6} + \frac{6}{6}$

$= 7/6$

$var(E[X|Y]) = var(X) - E[var(X|Y)]$

(or) straight forward

$var(E[X|Y]) = \frac{1}{3}\left(\frac{1}{2} - \frac{7}{6}\right)^2 + \frac{2}{3}\left(\frac{3}{2} - \frac{7}{6}\right)^2$

$= \frac{1}{3}\left(\frac{4}{9}\right) + \frac{2}{3}\left(\frac{1}{9}\right)$

$= \frac{4}{27} + \frac{2}{27} = \frac{6}{27} = \frac{2}{9}$

$E[var(X|Y)] = \frac{1}{3}\left(\frac{1}{12}\right) + \frac{2}{3}\left(\frac{1}{2}\right)$

$= \frac{1}{36} + \frac{2}{36} = \frac{3}{36} = \frac{1}{12}$

$var(X) = \frac{2}{9} + \frac{1}{12} = \frac{11}{36}$

Sum of a R.V number of independent R.V's

N: number of stores visited (Non negative integer r.v)

X_i: money spent in each store

→ X_i – assumed iid (independent & identically distributed)

* Independent of N

* Each store: Fresh men

$y = X_1 + X_2 + \cdots X_N$ → (how many R.v is not gn)

N=n, n number of stores

$X_1 + \cdots X_n$ money spent.

$E[Y|N=n] = E[X_1 + X_2 + \cdots X_n \mid N=n]$

$= E[X_1 + X_2 + \cdots + X_n]$ → ∴ Independent.

$= E[X_1] + E[X_2] + \cdots + E[X_n]$

Expectation of Sum = Sum of Expectation
↓
fixed no. of R.V

$= n\,E[X]$
↳ Expectation in a typical store.

* $E[Y|N] = N E[X]$ → Random variable.

$E[Y] = E[E[Y|N]]$

$= E[N E[X]]$

$= E[N]\,E[X]$

∴ E[X] is a number (constant)

∴ Inference: Expected amount to be spent in each store & Expected number of stores going to visit.

$Var(Y) = E[Var(Y|N)] + Var[E[Y|N]]$

$E[Y|N] = N E[X]$

* $Var(E[Y|N]) = Var(N E[X]) = (E[X])^2\,Var(N)$

* $Var(Y|N=n) = n\,Var(X)$ [Each store has a variance
↓
n stores.]

$Var(Y|N) = N\,Var(X)$

$E[Var(Y|N)] = E[N]\,Var(X)$

Var(X) → Constant

∴ $Var(Y|N=n) = Var(X_1 + X_2 + \cdots X_n)$

$= Var(X_1) + \cdots + Var(X_n)$

$= n\,Var(X)$
↳ of a typical store

$Var[Y] = E[Var(Y|N)] + Var(E[Y|N])$

$$\text{var}[Y] = E[N]\,\text{var}\,X + (E[X])^2\,\text{var}(N)$$

var[Y] ↓ Total amount

E[N] var X ↓ variability how much spend in each store

(E[X])² var(N) ↓ variability in visiting each store.

Recitation

Stick breaking problem:

$$E[X] = E[E[X|Y]] \xrightarrow{} \text{function of } y$$

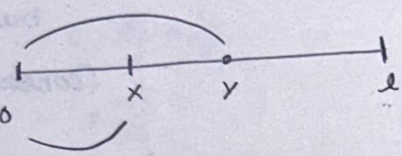
Law of total variation:

$$\text{var}(X) = E[\text{var}(X|Y)] + \text{var}(E[X|Y])$$

if $Y \sim U[a,b]$, $\text{var}(Y) = \frac{(b-a)^2}{2}$

↓ uniform distribution

$$E[Y] = \frac{a+b}{2}$$



a) $Y \sim U[0,l]$

E[length remaining after breaking twice]

variance (")

$$E[Y] = \frac{l}{2}$$

$$\text{var}(Y) = \frac{l^2}{12}$$

$X \sim [0, Y] \rightarrow$ Random variable.

Given, $Y = y$

$$X \sim U[0, y]$$

$$E[X|Y] = \frac{Y}{2}$$

$$\text{var}(X|Y) = \frac{Y^2}{12} \rightarrow \text{R.V}$$

$$E[X] = E[E[X|Y]]$$

$$= E\left[\frac{Y}{2}\right] = \frac{1}{2}E[Y]$$

$$\boxed{E[X] = \frac{l}{4}.}$$

b) $\text{var}(X) = E[\text{var}(X|Y)] + \text{var}(E[X|Y])$

$$E[\text{var}(X|Y)] = E\left[\frac{Y^2}{12}\right]$$

$$= \frac{1}{12} E[Y^2]$$

$$\text{var}(X) = E[X^2] - (E[X])^2$$

$$= \frac{1}{12}\left(\text{var}(Y) + (E[Y])^2\right)$$

$$= \frac{1}{12}\left(\frac{l^2}{12} + \left(\frac{l}{2}\right)^2\right)$$

$$= \frac{l^2}{36}$$

$$\text{var}(E[X|Y]) = \text{var}\left(\frac{Y}{2}\right)$$

$$= \frac{1}{4}\text{var}(Y)$$

$$= \frac{1}{4}\frac{l^2}{12}$$

$$= \frac{l^2}{48}$$

$$\boxed{\text{var}(X) = \frac{l^2}{36} + \frac{l^2}{48} = \frac{7l^2}{144}}$$

---

widgets & crates

N: # of boxes

X_i: # widgets of a box

T: total # of widgets

$E[X_i] = E[N] = 10$

$\text{var}(X_i) = \text{var}(N) = 16$

X_i's, N independent.

$T = X_1 + X_2 + \dots + X_N$

$S = X_1 + X_2 + \dots + X_{12}$ → Total number of boxes (Assume)

Linearity:

$E[S] = 12\,E[X_i]$

$\text{var}(S) = 12\,\text{var}(X_i)$

$X_1, X_2 \dots$ → widgets.

conditioning on N

$E[E[T|N]] = E[T]$

→ Total widgets = widgets in each

$E[T] = E[E[T|N]]$

$= E[N\,E[X_i]]$

$$\boxed{E[T] = E[X_i]\,E[N]}$$

$E[T] = 10\,E[N]$

$E[T] = 100$

∴ 10 boxes, 10 widgets each

b) If the distribution depends on (no. of widges) boxes → our model is no longer valid.

→ Total variance = variance in each.

$$Var(T) = E[Var(T|N)] + Var(E[T|N])$$

$$= E[N\, Var(X_i)] + Var(N\, E[X_i])$$

$$= E[N \cdot 16] + Var(N \cdot 10)$$

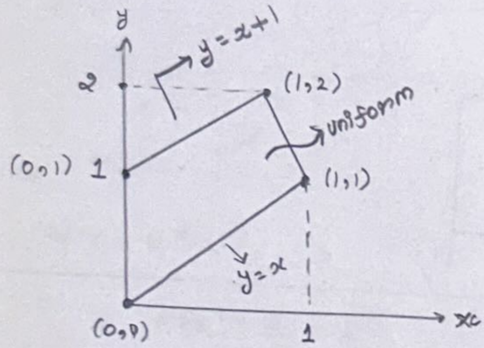
$$= 16\, E[N] + 100\, Var(N)$$

$$= 16 \times 10 + 100 \times 16$$

$$= 160 + 1600$$

$$= 1760.$$

Using the cond. Expectation & variances



x, y are uniformly distributed (Area = 1)

$$Var(x+y) = Var(E[x+y|x]) + E[Var(x+y|x)]$$

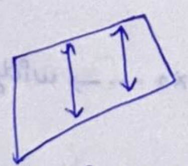
$$= Var(E[x+y|y]) + E[Var(x+y|y)]$$

↓

we can condition on something

Condition on particular x:

width stays constant

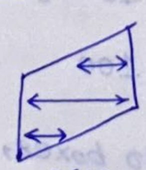


→ So condition on x.

• Same width

Condition on y:

$E[x+y|x] = ?$

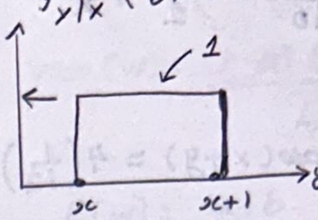


Different width

Conditioning doesn't change relative frequency of outcome.

y varies from x to $x+1$

$\therefore$ x is uniformly distributed $\rightarrow$ y also.



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

$$E[X+Y|X] = E[X|X+Y|X]$$

$\downarrow$
constant
[$\because$ x conditioned on x]

$\therefore P_{Y|X}(y|x)$

$\therefore$ knowing x, knowing y.

$$= E[X+Y|X]$$

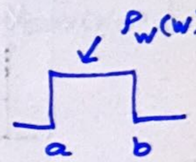
$$= E[X] + E[Y|X]$$

$$\therefore E[Y|X] = \frac{x+1+x}{2} = \frac{2x+1}{2}$$

$$= X + E[Y|X]$$

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

$$= \frac{4x+1}{2}$$



$$\text{var}(W) = \frac{(b-a)^2}{12}$$

$$E[\text{var}(X+Y|X)] = E[\text{var}(X|X+Y|X)]$$

$$= \text{var}(X) + \text{var}(Y|X)$$

$$= 0 + \text{var}(Y|X)$$

$$\text{var}(Y|X) = \frac{(x+1-x)^2}{12} = \frac{1}{12}$$

$$= 0 + \frac{1}{12}$$

$$= 1/12$$

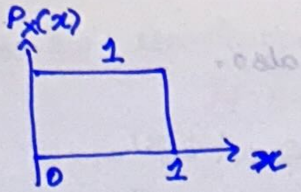
$$\text{var}\left(2X+\frac{1}{2}\right) = \text{var}(2X) = 4\,\text{var}(X)$$

$$\text{var}(X+Y) = \text{var}\left(2X+\frac{1}{2}\right) + E\left[\frac{1}{12}\right]$$

$$= 4\,\text{var}(X) + \frac{1}{12}$$

$$\text{var}(X) = ?$$

• width of $\boxed{\updownarrow\updownarrow\updownarrow}$ is same $\rightarrow$ variance is same



$var(X) = \frac{(1-0)^2}{12} = \frac{1}{12}$

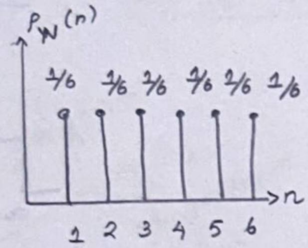
$var(X+Y) = 4\left(\frac{1}{12}\right) + \frac{1}{12} = \frac{5}{12}$

A random no. of coin flips

* Fair 6 sided die
* Fair coin — # times flipped (number of die)
* want expectation & variance of # heads

a) Repeat case a → For a dice (instead of 1)

Solu::



H → # heads

N → # out of die roll (no. of coin flips)

$$X_i = \begin{cases} 1 & \text{, heads on } i \text{ th flip} \\ 0 & \text{, Tails on } i \text{ th head} \end{cases}$$

$H = X_1 + X_2 + X_3 + \cdots + X_N$

$\boxed{\therefore \text{ If head } X_i = 1}$

$E[H] = E[X_1 + X_2 + \cdots + X_N]$ — Random no. of random variable.

$= E[X_1] + \cdots + E[X_N]$

~~[we don't able to say~~ like: ~~money spent in a shop~~]

~~Binary(...)~~

can't be done

Previous problems: $E[\;] = N\,E[X_i]$

∴ N is a random variable too.

∴

$\therefore\ E[H] = E[E[H/N]] = E[E[X_1 + \cdots X_n]]$

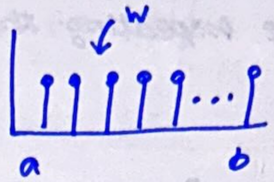
$= E[N\,E[X_i]]$

$= E[X_i]\ E[N].$

$= E[N]\left[\frac{1}{2}(1) + \frac{1}{2}(0)\right]$

$= E[N] \cdot \frac{1}{2}$ what about $E[N]$

In discrete case



$$var(w) = \frac{(b-a)(b-a+2)}{12}$$

$$E[w] = \frac{a+b}{2}$$

$\therefore N \to$ numbers in die (6)

$$E[H] = \frac{1}{2} E[N]$$
$$= \frac{1}{2}\left(\frac{1+6}{2}\right)$$
$$= \frac{7}{4}$$

$$var(H) = var(E[H|N]) + E[var(H|N)]$$
$$= var(E[x_1 + \dots + x_N | N]) + E[var(x_1 + \dots + x_N | N)]$$
$$= var(E[x_1] + E[x_2] + \dots E[x_N]) + E[var(x_1) + \dots + var(x_n)]$$
$$= var(N E[x_i]) + E[var(x_i) \cdot N]$$
$$= (E[x_i])^2 var(N) + var(x_i) E[N]$$
$$= \left(\frac{1}{2}\right)^2 \left(\frac{(6-1)(6-1+2)}{12}\right) + var(x_i)\left(\frac{1+6}{2}\right)$$
$$= \frac{1}{4}\left(\frac{5 \times 7}{12}\right) + \left(\frac{1}{4}\right)\left(\frac{7}{2}\right)$$
$$= \frac{77}{48}$$

$x_i \to$ Bernoulli R.V

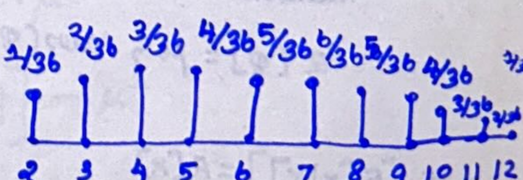
$$var(x_i) = p(1-p)$$
$$= \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4}$$

2 dice

$N_1 \to$ outcome of 1st die

$N_2 \to$ out come of second die.

coinflips $= N_1 + N_2$



$H_1 \to$ # heads in N_1 flips

$H_2 \to$ # heads in N_2 flips.

'Binomial pmf'

so splitting: Twice the first case (1 time)

↓

we are repeating the experiment twice.

$\boxed{H^* = H_1 + H_2}$

h_1 & h_2 are independent. [All of our coin flips are experiment]

$\rightarrow E[H^*] = E[H_1] + E[H_2]$

$= \frac{7}{4} + \frac{7}{4}$

$= \frac{14}{4} = \frac{7}{2}$

$var(H^*) = var(H_1) + var(H_2)$

$= 2\left(\frac{77}{48}\right)$

$= \frac{77}{24}$.

A coin with a random bias

prob (head) = v is the value of r.v. q with gn mean μ & positive variance σ^2. Let X_i be a Bernoulli r.v — models outcome of ith toss ($X_i = 1$ if the i th toss is head.

$X_1, X_2 \dots, X_n$ are conditionally independent,

Gn: $Q = v$, Let X be the no. of heads in n tosses

i) use the law of iterated expectations to find $E[X_i]$ and $E[X]$

ii) $Cov(X_i, X_j)$ Are X_i, X_j indp.

iii) use total variance — $Var(X)$. verify covariance result

Solu:

Random bias Q

$E[Q] = \mu$, $Var(Q) = \sigma^2 > 0$

Gn
$X_1, X_2, \dots, X_n$ are
Cond. independent

$\star\ E[E[X|Y]] = E[X]$

$\star\ Cov(X, Y) = E[XY] - E[X]E[Y]$

$\star\ var(X) = E[var(X|Y)] + var(E[X|Y])$

$X \sim Bernoulli(P)$, $E[X] = P$, $Var(X) = P(1-P)$

a) $E[x_i] = E[E[x_i \mid \varphi]]$

$= E[E[x_1+x_2+\cdots+x_n \mid \varphi]]$

$= E[E[x \mid \varphi]]$

$= E[\varphi]$
↓
R.V

$= E[\varphi]$

$\boxed{E[x_i] = \mu}$

$x_i = 1 \rightarrow$ head

$x_i = 0 \rightarrow$ tail

$x_1 + \cdots + x_n = X$ (Number of heads)

No. of heads = α [Bias]

$\therefore E[X] = P$
↓
gn bias

[our case we don't know α $\varphi \rightarrow$ R.V]

a) $E[X] = E[x_1 + \cdots + x_n]$

$(x_1+\cdots x_n) = E[$ ~~...~~

$= E[E[x_1+\cdots+x_n \mid \varphi]]$

$= E[n\, E[x_i]]$

$= n\, E[x_i]$

$= n\mu$

$x_i \rightarrow$ particular toss.

$X \rightarrow$ total tosses

(n tosses)
↓
Previous cases we don't know n (So R.V = N)
↓
This case n is known
Bias φ is unknown
↓
R.V

b) $\mathrm{cov}(x_i, x_j) = E[XY] - E[X]\,E[Y]$

↙ i, j same — Same toss

↘ i, j different — different toss.

① $i \neq j$

$\mathrm{cov}(x_i, x_j) = E[x_i x_j] - E[x_i]\,E[x_j]$

$= E[E[x_i x_j \mid \varphi]] - \mu\cdot\mu$

x_i, x_j are conditionally independent

$= E[E[x_i \mid \varphi]\, E[x_j \mid \varphi]] - \mu^2$

$= E[\varphi\cdot\varphi] - \mu^2$

$= E[\varphi^2] - \mu^2$

$= \mathrm{var}(\varphi) = \sigma^2 > 0$

i. As $\mathrm{cov}(x_i, x_j) \neq 0$

Correlative

They can't be Independent.

② when $i = j$

$$cov(x_i, x_i) = E[x_i^2] - (E[x_i])^2$$

$$= E[E[x_i^2 | Q]] - (E[x_i])^2$$

x_i is the same

Bernoulli $\rightarrow 1 \rightarrow 1^2 = 1$
$\rightarrow 0 \rightarrow 0^2 = 0$

$$= E[Q] - \mu^2 \qquad = var(x_i)$$

$$\boxed{var(x_i) = \mu - \mu^2} \neq 0 \qquad \boxed{\therefore cov(x_i, x_j) = var(x_i)}$$

' x_i & x_j are not independent '

$\star\ E[x_i] = \mu,\ cov(x_i, x_j) = \sigma^2\ (i \neq j)$

$$\boxed{var(x_i) = \mu - \mu^2}$$

c) Total variance

$x_1, x_n .. \rightarrow$ condt independent.

$$var(x) = E[var(x|Q)] + var(E[x|Q])$$

$$E[var(x|Q)] = E[var(x_1 + .. + x_n | Q)]$$

$$= E[var(x_1|Q) + var(x_2|Q) + \dots var(x_n|Q)]$$

$$= E[n\ var(x_i|Q)]$$

$$= E[n\ var(x_i|Q)]$$

$x_i \rightarrow$ Bernoulli

For Bernoulli

$var(x_i|Q) = Q(1-Q)$

Bias $\rightarrow$ unknown (R.V)

$$= E[n\ Q(1-Q)]$$

$$= n\ E[Q - Q^2]$$

$$= n(E[Q] - E[Q^2])$$

$$= n(\mu - (var(Q) + (E[Q])^2))$$

$$= n(\mu - (\sigma^2 + \mu^2))$$

$$= n(\mu - \sigma^2 - \mu^2)$$