∵ the flips of each coin are independent
$P(A,B) = P(A)\,P(B)$
↳ this need not always be true!

L5

Chain rule of probability for finding jt. D^n.

$P(X=x, Y=y) = P(Y=y \mid X=x)\,\underbrace{P(X=x)}_{\text{marginal } D^n}$ → this follows from defn. of conditional prob.

$P(A|B) = \frac{P(A \cap B)}{P(B)}$

where $\sum_i \sum_j P(X=x_i, Y=y_j) = 1$

The generalization to n-variable case :-

$P(X_1 = x_1, \ldots, X_n = x_n)$

$= P(X_1 = x_1) \times P(X_2 = x_2 \mid X_1 = x_1) \times P(X_3 = x_3 \mid X_1 = x_1, X_2 = x_2)$

$\times \ldots \times P(X_n = x_n \mid X_1 = x_1, \ldots, X_{n-1} = x_{n-1})$.

Similarly in the continuous case

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

where $f_X(x) = \int_y f_{XY}(x,y)\,dy$ & $f_Y(y)$ } are marginal D^n.

$\iint_{x\,y} f_{XY}(x,y) = 1$

eg: Multi (Bi) - variate Normal D^n (Continuous D^n) is the most commonly encountered D^n in Statistics.

Scanned by CamScanner

L5

Jt. Cumulative D^n

$$F(x,y) = P(X \le x, Y \le y) = \sum_{s \le x} \sum_{t \le y} f_{XY}(s,t)$$

eg of jt/marginal D^n for discrete case -

$f(x,y) = P(X=x, Y=y)$ X = word length, Y = no. of vowels

from a survey

eg $f(2,1) = P(\text{to}) + P(\text{of}) + P(\text{on}) = 0.18 + 0.10 + 0.06 = 0.34$.

$f(3,0) = P(\text{BBC}) = 0.03$

$f(4,3) = 0$.

full D^n $f(x,y)$

y \ X	2	3	4	5	$\sum_x f(x,y) = f_Y(y)$
0	0	0.03	0	0	0.03
1	0.34	0.30	0.16	0	0.80
2	0	0	0.03	0.14	0.17
$\sum_y f(x,y) = f_X(x)$	0.34	0.33	0.19	0.14	

$$\begin{array}{r} 0.34 \\ \times 0.03 \\ \hline 0.0102 \end{array}$$

Note $\sum_x f_X(x) = \sum_y f_Y(y) = \sum_{x,y} f(x,y) = 1$

$f_X(x) f_Y(y)$

y \ X	2	3	4	5
0	$0.34 \times 0.03 = 0.01$	0.01	0.0057	0.0042
1	0.272	0.264	0.152	0.112
2	0.0578	0.0561	0.0323	0.0238

Clearly $f_X(x) f_Y(y) \ne f_{XY}(x,y) \Rightarrow$ X & Y are NOT indep.

Reading Assignment :- Practise some problems for jt/Marginal D^n for continuous Case.

(L5)

(13) D^n for transformations of Continuous RVs

Let $f_X(x)$ is known

We need to find the D^n of $Y = g(X)$

Let g be invertible $\Rightarrow$ g^{-1} is Strictly inc/dec!

Case (I) :- g^{-1} is increasing

$$F_Y(y) = P(Y \le y)$$
$$= P(g(X) \le y)$$
$$= P(X \le g^{-1}(y))$$
$$= F_X(g^{-1}(y))$$

(ineq is preserved b/c g^{-1} is inc.)

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$
$$= \frac{d}{dy} F_X(g^{-1}(y)) \stackrel{\text{chain rule}}{=} \frac{d}{dx} F_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

Case (II) :- g^{-1} is dec.

$$F_Y(y) = P(Y \le y) = P(g(X) \le y)$$
$$= P(X \ge g^{-1}(y))$$

(b/c g^{-1} is dec.)

$$= 1 - P(X < g^{-1}(y))$$
$$\stackrel{X \text{ is Cont RV}}{=} 1 - P(X \le g^{-1}(y))$$
$$= 1 - F_X(g^{-1}(y)).$$

Scanned by CamScanner

L5

14 Some examples of D^n of composite RVs.

14.1) Let X & Y be independent geom$_1(p)$

(a) find the D^n of $\min(X,Y)$

(b) $P(Y \geq X)$

(c) Find D^n of $X+Y$

(d) $P(Y=y \mid \underbrace{X+Y}_{Z} = z)$ for $z \geq 2$; $y = 1, \dots, z-1$

Soln: (a) $P(\min(X,Y) > z)$ i.e. $P(Z > z)$

$= P(X \geq z, Y > z)$

$\stackrel{\text{indep}}{=} P(X > z)P(Y > z)$

$= [1 - P(X \leq z)][1 - P(Y \leq z)]$

$= \left[1 - \sum_{i=1}^{z} P(X = i)\right]\left[1 - \sum_{j=1}^{z} P(Y = j)\right]$

$= \left[1 - \sum_{i=1}^{z} (1-p)^{i-1} p\right]\left[1 - \sum_{j=1}^{z} (1-p)^{j-1} p\right]$

$\stackrel{\text{geom series sum}}{=} \left[1 - (1-(1-p)^{z-1})\right]\left[1 - (1-(1-p)^{z-1})\right]$

$= (1-p)^{z-1}(1-p)^{z-1} = \left[(1-p)^2\right]^{(z-1)}$

$P(Z > z) = (1-p)^{2(z-1)}$

i.e. $\min(X,Y) \sim$ geom$_1(1-(1-p)^2)$ (See next pg).

$\cancel{\therefore P(\min(X,Y) \leq z) = 1 - (1-p)^{2(z-1)}}$

Note if $X \sim$ geom$_1(p)$; $f_X(x) = (1-p)^{x-1} p$.

$P(X > x) = 1 - P(X \leq x) = 1 - \sum_{m=1}^{x} P(X = m)$

$= 1 - \sum_{m=1}^{x} (1-p)^{m-1} p$

$= 1 - p\sum_{m=0}^{x-1} (1-p)^{m}$

$$P(X>x) = 1 - p\left\{\frac{1-(1-p)^x}{1-(1-p)}\right\}$$

$$= 1 - 1 + (1-p)^x$$

$$= (1-p)^x$$

$$P(X \le x) = 1 - (1-p)^x \checkmark$$

L5

Also note:-

$$P(X \ge x)$$
$$= P(X=x) + P(X>x)$$
$$= (1-p)^{x-1}p + (1-p)^x$$
$$= (1-p)^{x-1}(p+1-p)$$
$$= (1-p)^{x-1}$$

$$1-q = (1-p)^2$$

$$\Rightarrow q = 1-(1-p)^2 \checkmark$$

(b) $$P(Y \ge X) = \sum_{x=1}^{\infty} P(X=x, Y \ge X)$$

$$= \sum_{x=1}^{\infty} P(X=x, Y \ge x)$$

$$\overset{\text{indep.}}{=} \sum_{x=1}^{\infty} P(X=x)P(Y \ge x)$$

$$= \sum_{x=1}^{\infty} (1-p)^{x-1}p\,(1-p)^{x-1} \checkmark$$

$$= p\sum_{x=1}^{\infty} (1-p)^{2(x-1)}$$

$$= \frac{p}{2p-p^2}$$

(c) Let $z \ge 2$ & integer

$$P(X+Y=z) = \sum_{x=1}^{z-1} P(X=x, X+Y=z)$$

$$= \sum_{x=1}^{z-1} P(X=x, Y=z-x)$$

$$\overset{\text{indep.}}{=} \sum_{x=1}^{z-1} P(X=x)P(Y=z-x)$$

$$= \sum_{x=1}^{z-1} p(1-p)^{x-1} p(1-p)^{z-x-1}$$

$$= (z-1)p^2(1-p)^{z-2}$$

L5

(d) $P(Y=y \mid X+Y=z)$

$$= \frac{P(Y=y, X+Y=z)}{P(X+Y=z)}$$

$$= \frac{P(X=z-y, Y=y)}{P(X+Y=z)} \overset{\text{indep}}{=} \frac{P(X=z-y)P(Y=y)}{P(X+Y=z)}$$

$$= \frac{p(1-p)^{z-y-1} p(1-p)^{y-1}}{(z-1)p^2(1-p)^{z-2}}$$

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

(15) <u>Sums of independent "continuous" RVs</u>

$$F_{X+Y}(z) = P(X+Y \le z)$$

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

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_X(u-y) f_Y(y)\,du\,dy$$

$$= \int_{-\infty}^{z} \int_{-\infty}^{\infty} f_X(u-y) f_Y(y)\,dy\,du$$

$$\& \therefore f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,dy = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx$$

Note :-

L5

Defn (Convolution)

$$(f * g)(z) = \int_{-\infty}^{\infty} f(z-y)\,g(y)\,dy = \int_{-\infty}^{\infty} g(z-x)\,f(x)\,dx$$

$$\therefore Z = X + Y \sim f_Z(z) = (f * g)(z).$$

indep

---

(15) **Useful Identity** :-

(1) $P(S < T) = \int_{s} f_S(s)\,P(T > s)\,ds$

This should be intuitively clear by following the preceeding solved problem for the discrete case!

(11) $E(I\{X \geq k\}) = 1\,P(X \geq k) + 0 \cdot P(X < k)$
$= P(X \geq k)$

So. $\boxed{P(X \geq k) = E(I\{X \geq k\})}$ This is often v. useful!

L6

16 Moment generating f^n. (mgf)

X is a R.V.

$$M_X(t) := E\left(e^{tX}\right)$$

Utility of mgf.

1. To compute moments of D^n
   i.e. $E(X), E(X^2), E(X^3) \ldots$ b/c

$$\boxed{\begin{aligned} E(X^k) &= \left.\frac{d^k}{dt^k} M_X(t)\right|_{t=0} \\ &= M_X^{(k)}(0) \end{aligned}}$$

2. $X_i \overset{iid}{\sim}$ R.D (rand. D^n).

$$Y = X_1 + X_2 + \cdots + X_n$$

$$\begin{aligned} M_Y(t) &= E\left(e^{t(X_1+X_2+\cdots+X_n)}\right) \\ &= E\left(e^{tX_1} e^{tX_2} e^{tX_3} \cdots e^{tX_n}\right) \\ &\overset{\text{indep}}{=} E\left(e^{tX_1}\right) E\left(e^{tX_2}\right) \cdots E\left(e^{tX_n}\right) \\ &= M_{X_1}(t) M_{X_2}(t) \cdots M_{X_n}(t) \\ &= \left(M_X(t)\right)^n \quad ; X \sim X_i \end{aligned}$$

& then try to identify D^n of Y.

b/c

$$\boxed{\text{a mgf uniquely determines the } D^n \text{ of a RV.}}$$

(17) Important statistical moments.

L6

1) 1st moment is the mean: $E(X)$.

2) 2nd moment is related to the variance:- $E(X^2) = var(X) + [E(X)]^2$

3) Normalized 3rd moment (skewness).

$$\gamma_1 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{E[(X-\mu)^3]}{\left(E[(X-\mu)^2]\right)^{3/2}} = \frac{K_3}{K_2^{3/2}}$$

Also called Pearson's moment coeff of skewness

$$= \frac{E(X^3) - 3\mu\sigma^2 + 2\mu^3}{\sigma^3}$$

Where $\mu_n = E[(X-E(X))^n]$ is the n^{th} central moment

Note $\mu_1 \neq \mu$ or $E(X)$

$\mu_1 \equiv 0$ b/c $E(X - E(X))$
$= E(X) - E(E(X))$
$= E(X) - E(X)$
$= 0$.

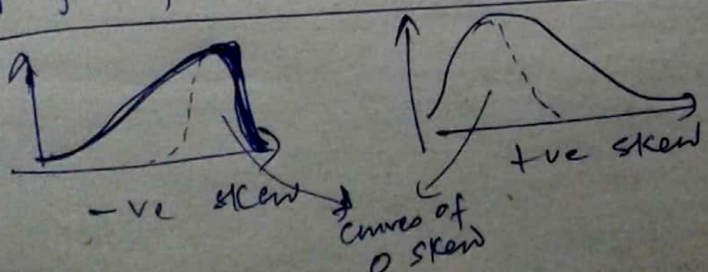
σ is std. dev.

The cumulant generating f^n

$$K(t) := \log\left(E(e^{tX})\right) = \sum_{n=1}^{\infty} K_n \frac{t^n}{n!} = \mu t + \sigma^2 \frac{t^2}{2} + \cdots$$

$K_n = K^{(n)}(0)$ (n^{th} derivative of $K(t)$ evaluated at $t=0$)

$$\boxed{\text{Skewness is a measure of asymmetry of pdf of a real valued RV about its mean.}}$$



γ_1 may be undefined!

Applications of skewness

turbulence signatures of velocity fluctuations

$v' = v - \langle v \rangle$

Skewness, $\gamma_1 = \dfrac{\langle v'^3 \rangle}{\langle v'^2 \rangle^{3/2}}$ where here $\langle x^k \rangle$ means $E\left((X - E(X))^k\right)$

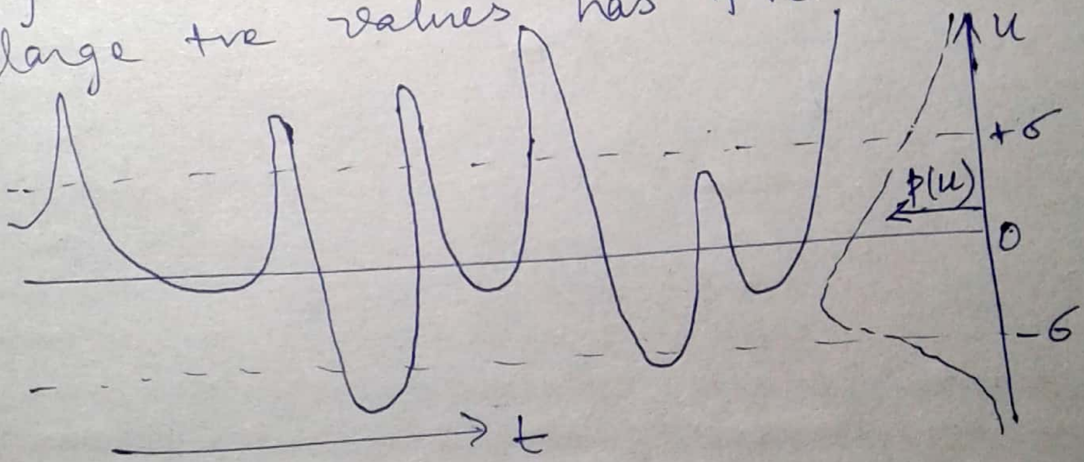
Note $\langle v' \rangle = 0$

so $\langle v'^k \rangle = E\left[(v' - \langle v' \rangle)^k\right]$

$= E(v'^k)$

+ve γ_1 means v' is more likely to take on large +ve values than large -ve values.

A time series signal w/ large stretches of small -ve values & few instances of large +ve values has +ve skewness.



4) Normalized 4th order moment - (Kurtosis/Flatness)

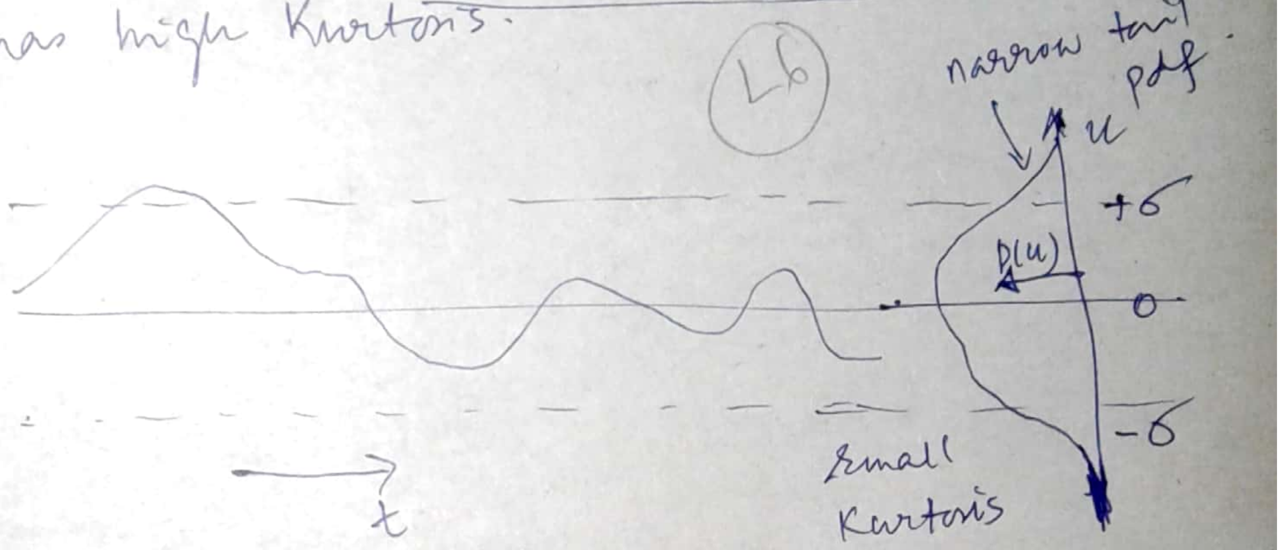
$$\mathrm{Kurt}(X) = E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right) = \frac{\mu_4}{\sigma^4} = \frac{E\left((X-\mu)^4\right)}{\left(E\left[(X-\mu)^2\right]\right)^2}$$

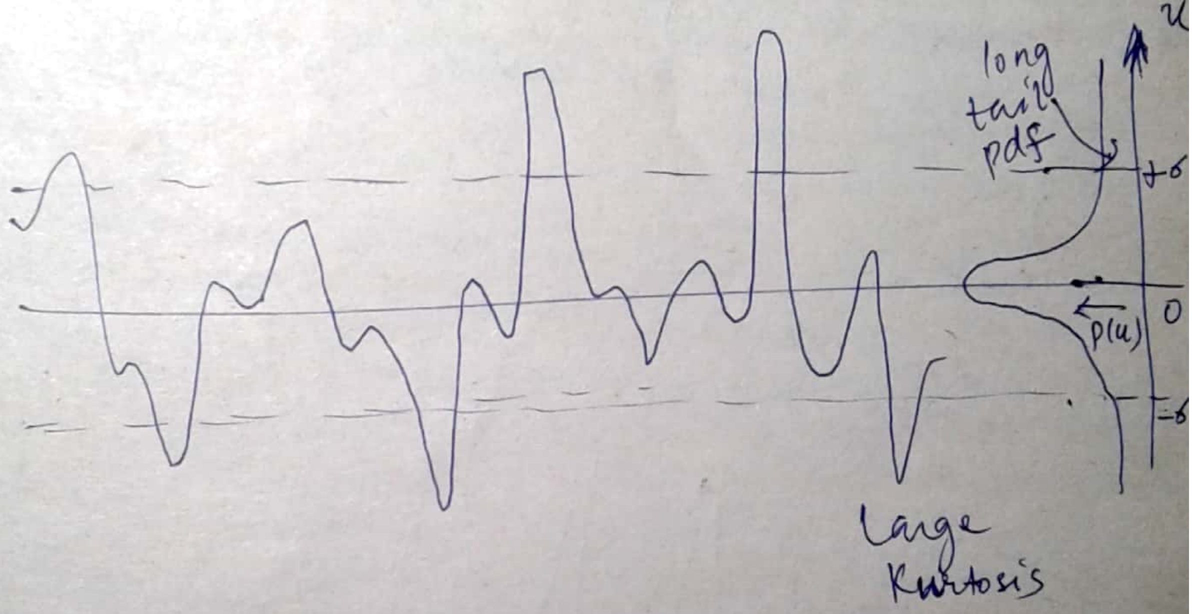
$$= \frac{\langle v'^4 \rangle}{\langle v'^2 \rangle^2}$$

* a pdf w/ longer tails will have a larger kurtosis than a pdf w/ narrower tails
* a time series w/ most measurements clustered around the mean has low kurtosis

* a time series dominated by intermittent events has high Kurtosis.

L6





(18) Law of large No.s & CLT

(L6)

(I) Markov's Inequality :

X is R.V. ; $r, c > 0$ (constant no.s)

$$P(|x| \geq c) \leq \frac{E(|x|^r)}{c^r}$$

(II) X is a R.V.

$g(x) > 0$ real valued f^n

for any $c > 0$

$$P(g(x) \geq c) \leq \frac{E(g(x))}{c}$$

(III) Chebyshev's inequality :

X is a RV w/ mean μ & variance $\sigma^2 < \infty$

then for any $k > 0$

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

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

(A) Weak Law of Large no.s

$$\overline{X}_n \xrightarrow{P} \mu \text{ as } n \to \infty$$

Def^n of conv. in probability { i.e. $\lim_{n \to \infty} P(|\bar{X}_n - \mu| > \epsilon) = 0$ for any $\epsilon > 0$

(B) Strong law of large no.s.

$$\overline{X}_n \xrightarrow{a.s} \mu \text{ as } n \to \infty$$

$$P(\lim_{n \to \infty} \bar{X}_n = \mu) = 1$$

Application :- eg, while a casino may lose money in a single spin of the roulette wheel, its earnings will tend towards a predictable % over a large no. of spins. Any winning streak by a player will eventually be overcome by the parameters of the game.

L6

* There are instances when the strong law does not hold but the weak law does.

(15) CLT (Application → when we study hypothesis testing).

Let $X_1, X_2, \ldots, X_n$ be a random sample from a D^n w/ mean μ & variance $\sigma^2 < \infty$

then $Z_n := \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \xrightarrow{D} Z \sim N(0,1)$ as $n \to \infty$.

(I) Conv. in probability.

$\{X_n\}$ is a seq. of RV

$X_n \xrightarrow{P} X$ if $\forall \epsilon > 0, \lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$

(II) Conv. in D^n

$\{X_n\}$ is a seq of RV w/ cdf $F_n(x)$

Let $F_n(x) = P(X_n \le x)$;

X is a RV w/ cdf $F(x) := P(X \le x)$

then $X_n \xrightarrow{D} X$ if $\lim_{n\to\infty} F_n(x) = F(x)$

Note.

$M_n(t) \xrightarrow{n\to\infty} M(t)$

$\Rightarrow X_n \xrightarrow{D} X$

* $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X$ (reverse not true)

except, $X_n \xrightarrow{D} c$ (const) $\Rightarrow X_n \xrightarrow{P} c$ (const).