Moment Generating Functions

Stochastic Processes. Lecture 01.

Independence of events

- Given multiple events defined on a sample space, **independence** of these events is true iff they are *pairwise* independent AND *collectively* independent.
- Pairwise independence does not imply collective independence and inverse is also not true.
- Expectations of jointly continuous random variables. Case when these are independent.
- Independence is a stronger criteria. Independence implies that covariance is zero. $E(XY) = E(X)E(Y)$. Inverse not true.
- Theorem. $Var(X + Y) = Var(X)Var(Y) + 2Cov(X, Y)$ proved.

Moment Generating functions

- Moment generating functions defined and connection to Taylor's series expansion shown.
- Example. Find $M_x(t)$ for the binomial random variable given parameters n and p.
- Discussion. If the moment generating functions are same for two RV's for all t, then they have the same means and variances. In fact, they have the same distribution.
- MGF of Poisson distribution and standard normal distribution derived.
- Properties of MGF. Convolution property.
- Theorem. If X and Y are normally distributed, so is the RV $Z = X + Y$.
- Building up to the central limit theorem.

Conditions for Independence of events

Sample space - S. ← set of all events.

independence: conditioning does not change prob of simultaneous occurence.

A_1, A_2, A_3 are 3 $\underbrace{\text{independent (mutually)}}_{\text{STRONG CONDITION}}$ events.

$A_j \subseteq S$.

C①
$$P(A_1A_2) = P(A_1)\,P(A_2)$$
$$P(A_2A_3) = P(A_2)\,P(A_3)$$
$$P(A_3A_1) = P(A_3)\,P(A_1)$$
} pairwise independence — weaker condition

C②
$$P(A_1A_2A_3) = P(A_1)\,P(A_2)\,P(A_3)$$
} collective independence. — stronger condition

Example 1. Toss a fair coin twice.

$S = \{HH, HT, TH, TT\}$.

$A_1 = \{HH, HT\}$ H in 1st toss

$A_2 = \{HH, TH\}$ H in 2nd toss

$A_3 = \{HH, TT\}$ same outcome in both tosses.

$$P(A_1A_2) = P(A_1 \cap A_2) = P(\{HH\}) = 1/4$$

$$P(A_1).\,P(A_2) = \frac{2}{4}\cdot\frac{2}{4} = 1/4 \quad \text{(uniform probability).}$$

$$\therefore \quad P(A_i\,A_j) = P(A_i)\,P(A_j) \quad \text{when } i \neq j$$

$$\left.\begin{array}{l} P(A_1A_2A_3) = P(\{HH\}) = 1/4 \\ P(A_1)P(A_2)P(A_3) = 2/4\cdot 2/4\cdot 2/4 = 1/8 \end{array}\right\} \neq$$

$\boxed{\therefore\ C① \not\Rightarrow C②}$

Example 2. Toss a fair die. $S = \{1, 2, 3, 4, 5, 6\}$.

$A_1 = \{1, 2, 3, 4\}$ $A_2 = \{4, 5, 6\}$ $A_3 = \{4, \cancel{5}, 6\}$.

$$\left.\begin{array}{l} P(A_1A_2A_3) = P(\{4\}) = 1/6 \\ P(A_1)P(A_2)P(A_3) = \frac{4}{6}\cdot\frac{3}{6}\cdot\frac{3}{6} = 1/6 \end{array}\right\} =$$

$\boxed{\therefore\ C② \not\Rightarrow C①}$

$$\left.\begin{array}{l} P(A_1)P(A_2) = \frac{4}{6}\cdot\frac{3}{6} = 1/3 \\ P(A_1A_2) = P(\{4\}) = 1/6 \end{array}\right\} \neq$$

From the axioms of probability, we know:

S : P (function)

If S is infinite there is some restriction in A.

"Measurability"

"Sigma algebra"

$P: P(S) \to \mathbb{R}$

↑ function.

If N is large, there are 2^N subsets for a set containing N elements.

Given any $A \subseteq S$, we have a value $P(A)$.

$P(S) \longrightarrow \mathbb{R}$

↑ all subsets of S

Real number line

Moment Generating Function:

↳ a nice differentiable function $f(x)$.

$f^n(x)$ exists for all n, infinitely differentiable.

$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$

$f'(x) = a_1 + 2a_2 x + \dots + n a_n x^{n-1} + \dots$

$f''(x) = 2a_2 + 3.2\, a_3 x + \dots \quad + n(n-1) a_n x^{n-2}$

$x = 0 \Rightarrow f(0) = a_0$

$f'(0) = a_1$

$f''(0) = 2a_2$

$f'''(0) = 3.2\, a_3$

$\vdots$

$f^{(n)}(0) = n!\, a_n \Rightarrow \boxed{a_n = \frac{1}{n!} f^{(n)}(0)}$

The power series expansion for $f(x)$ is unique and,

$$f(x) = f(0) + f'(0).x + \frac{f''(0)\, x^2}{2!} + \frac{f'''(0)\, x^3}{3!} + \dots$$

{Taylor's series}

X and Y are discrete random variables on which a PMF (probability mass function) is defined. → $P(x, y)$.

① $E(X+Y) = \sum_x \sum_y (x+y)\, P(x, y)$

$= \sum_x \sum_y x\, P(x, y) + \sum_x \sum_y y\, P(x, y).$

$= \sum_x x \underbrace{\sum_y P(x,y)}_{p(x)} + \sum_y y \underbrace{\sum_x P(x,y)}_{P(y)}$

$= \sum_x x\, P(x) + \sum_y y\, P(y)$

$= E(X) + E(Y).$

x is a constant in this summation.

$\sum_x \sum_y x\, p(x,y) = \sum_x \left[\sum_y x P(x,y) \right] = \sum_x x \underbrace{\left[\sum_y P(x,y) \right]}$

$= \sum_x x\, P(x) = E(X)$

$P(x) = \sum_y P(x,y)$ marginal dist. of R.V 'x'

{since probabilities are normalized values}

∴ $E(X+Y) = E(X) + E(Y)$ holds for discrete RVs.

If $E(X + X^2)$: we can calculate this, but not joint cases

In case of continuous random variables, if X, Y are jointly continuous, then $E(X+Y) = E(X) + E(Y)$

Analogously, this can be extended for a finite number of Random variables:

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E(X_i)$$ {LINEARITY of EXPECTATIONS OPERATOR}

Similar results exist for infinite random variables, but in such cases, one has to worry about CONVERGENCE.

② If X & Y are independent, then $E(XY) = E(X)\, E(Y)$
⇒ $\boxed{E(XY) = E(X)\, E(Y)}$

If X & Y are jointly discrete, with pmf $P(x, y)$.

$E(X, Y) = \sum_x \sum_y xy\, P(x, y)$

$$E(XY) = \sum_x \sum_y xy\, P(x, y)$$ X, Y are jointly discrete.

But X & Y are independent

$$\therefore P(x, y) = P_{(X,Y)}(x, y) = P_X(x)\, P_Y(y) \quad \forall\, x, y$$

as X & Y are independent

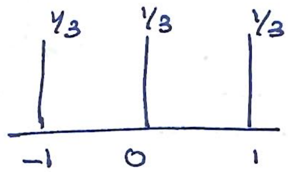
product of marginals.

So, $$E(X \cdot Y) = \underbrace{\sum_{x,y}}_{\text{double sum}} x\, y\, P_x(x)\, P_y(y) = \sum_x \left[\sum_y xy\, P_x(x)\, P_y(y) \right]$$

$$E(XY) = \left[\sum_x x\, P_X(x) \right] \left[\sum_y y\, P_Y(y) \right] = E(X)\, E(Y)$$

For independent RV's (discrete or cont.) $$E(XY) = E(X)\, E(Y)$$

Example.



PMF

$X = -1, 0, 1$

$$E[X] = (-1)\left(\tfrac{1}{3}\right) + 0\left(\tfrac{1}{3}\right) + 1\left(\tfrac{1}{3}\right) = 0$$

Let $Y = \begin{cases} 1 & \text{if } X = 0 \\ 0 & \text{otherwise} \end{cases}$

Defined in such a way that Y is dependent on X.

$$E(XY) = \underbrace{E(X)}_{=0} E(Y)$$

LHS = ? $\therefore$ RHS $= 0$

when $X = 0$, $Y \neq 0$, $XY = 0$
when $X \neq 0$, $Y = 0$, $XY = 0$ $\Rightarrow$ $XY = 0$ always $\quad E(XY) = 0$

Since $E(XY) = E(X)\, E(Y) = 0$, this expression holds,
But this does not mean X, Y are independent
X, Y are DEPENDENT here.

$$P(X = 1, Y = 1) \stackrel{?}{=} P(X = 1)\, P(Y = 1) \qquad 0 \neq \tfrac{1}{3} \cdot \tfrac{1}{3}$$

"," → and. $P(X=1, Y=1)$ → By defn '0', $P(X=1)$ → 1/3, $P(Y=1)$ → 1/3

$\therefore P(X=1, Y=1) \neq P(X=1) \cdot P(Y=1)$

Definition of independence:
$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B) \quad \forall\, A \,\&\, B.$$

INDEPENDENCE of X, Y $\Longrightarrow$ $E(XY) = E(X)\, E(Y)$, $\text{cov}(X, Y) = 0$

$E(XY) = E(X)\, E(Y) \not\Longrightarrow$ Independence.

③ $Var(X+Y) = E[(X+Y)^2] - [E(X+Y)]^2$

$$= E[X^2 + 2XY + Y^2] - [E(X) + E(Y)]^2$$

$$= E[X^2] + E[2XY] + E[Y^2] - (E(X))^2 \mp (E(Y))^2 - 2E(X)E(Y)$$

$$= \underbrace{E[X^2] - (E[X])^2}_{Var(X)} + \underbrace{E[Y^2] - (E[Y])^2}_{Var(Y)} + 2\underbrace{[E(XY) - E(X)E(Y)]}_{Cov(X, Y)}$$

$$\therefore \boxed{Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X, Y)}$$

Theorem: If X & Y are independent, then,

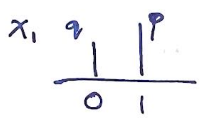
$Var(X+Y) = Var(X) + Var(Y)$ since $E(XY) = E(X)E(Y)$.

Example: n independent, identical trials and each with a prob. p of success. RV, $X = \#$ of successes.

$$X_j = \begin{cases} 1 & \text{if success on trial } j \\ 0 & \text{otherwise} \end{cases} \qquad j = 1, 2, 3, \ldots n$$

$$\therefore \quad X = \#\text{ of successes} = \sum_{j=1}^{n} X_j$$

Rigorous Proof.



$E(X_1) = q \cdot 0 + p \cdot 1 = p$

$Var(X_1) = E(X_1^2) - (E(X))^2 = (q \cdot 0^2 + p \cdot 1^2) - p^2$

$Var(X_1) = pq \qquad = p - p^2 = p(1-p) = pq$

But if the X_j's are independent,

$E(X) = \sum_{j=1}^{n} E(X_j) = np$ { Because of the linearity of expectation operator, if A, B are independent. $E(A+B) = E(A) + E(B)$. }

$Var(X) = \sum_{j=1}^{n} Var(X_j) = npq = \underbrace{Var(X_1)}_{pq} + \cdots + \underbrace{Var(X_n)}_{pq} = npq$.

Theorem: $Cov\left(\sum X_i, \sum Y_j\right) = \sum_i \sum_j Cov(X_i, Y_j)$

Moment-generating function

If X is a continuous or discrete random variables

then, $M_X(t) := E(e^{tX})$ ← Moment generating function MGF of RV. X

$E(X) :=$ first moment $=$ mean

$E(X^2) :=$ second moment of X

$E(X^k) := k^{th}$ moment of X.

$$M_X(t) = \begin{cases} \sum_x e^{tx} \underbrace{P(X=x)}_{\text{PMF}} & \text{if } X \text{ is a discrete RV.} \\ \int_{-\infty}^{\infty} e^{tx} \underbrace{f_X(x)}_{\text{prob. density function}} dx & \text{if } X \text{ is a continuous RV.} \end{cases}$$

Note: the MGF is a real valued function

$M_X(t) = E(e^{tX})$

$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ {Taylor series expansion for the exponential function}

$$e^{tX} = 1 + tX + \frac{(tX)^2}{2!} + \cdots + \frac{(tX)^k}{k!} + \cdots$$

$$e^{tX} = 1 + tX + \frac{t^2}{2!}X^2 + \cdots + \frac{t^k}{k!}X^k + \cdots$$

$$\underbrace{E[e^{tX}]}_{\text{MGF of } X} = \underbrace{E[1]}_{=1} + t\underbrace{E[X]}_{1^{st}\text{ moment}} + \frac{t^2}{2!}\underbrace{E[X^2]}_{2^{nd}\text{ moment}} + \cdots + \frac{t^k}{k!}\underbrace{E[X^k]}_{\text{Higher order moments}} + \cdots$$

∴ Diff MGF once and setting $t=0$ gives mean

Diff MGF twice & setting $t=0$ gives 2^{nd} moment & so on..

$$\Rightarrow M_X^{(k)}(0) = E(X^k)$$

Note: If X & Y are RV's.

If $M_X(t) = M_Y(t) \ \forall t$,

then X & Y have the same distribution.

Example: X is Binomial with parameter n & p.
Find $M_X(t)$. (X is a discrete RV).

$$M_X(t) = \sum_{k=0}^{n} e^{tk}\left(P(X=k)\right) = \sum_{k=0}^{n} e^{tk}\binom{n}{k} p^k q^{n-k}$$

$$M_X(t) = \sum_{k=0}^{n} \binom{n}{k}(e^t p)^k q^{n-k} = (e^t p + q)^n$$

$$M_X(t) = (e^t p + q)^n \, ;$$

$$M_X(t) = 1 + E(X)\,t + \frac{E(X^2)}{2!}t^2 + \cdots + \frac{E(X^k)}{k!}t^k + \cdots$$

and $M_X^{(k)}(0) = E(X^k)$

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n + \cdots \qquad a_n = \frac{f^{n}(0)}{n!}$$

$$M_X(t) = (e^t p + q)^n$$

$$M_X'(t) = n(e^t p + q)^{n-1} p e^t$$

$$M_X''(t) = n(n-1)(e^t p + q)^{n-2}(pe^t)(pe^t) + n(e^t p + q)^{n-1} p e^t$$

$$M_X'(0) = n(\underbrace{p+q}_{p+q=1})^{n-1} p = np = E(X)$$ = mean of the Binomial distribution.

$$M_X''(0) = n(n-1)(p+q)^{n-2}p^2 + n(p+q)^{n-1}p$$
$$= n(n-1)p^2 + np = n^2p^2 - np^2 + np = E(X^2).$$

$$Var(X) = \sigma^2 = E(X^2) - (E(X))^2 = n^2p^2 - np^2 + np - n^2p^2$$

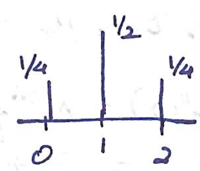
$$Var(X) = -np^2 + np = np(1-p) = npq = Var(X)$$
= variance of the Binomial distribution

Fact. If X, Y are RV's.
If $M_X(t) = M_Y(t)\ \forall t$, then X & Y have the same distribution

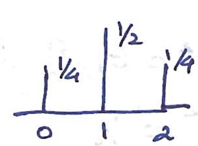
Toss a coin twice $X = \#$ heads, $Y = \#$ tails.
$X \neq Y$

Since sample space outcome HH

HH $\Rightarrow X = 2$
HH $\Rightarrow Y = 0$



pmf of X



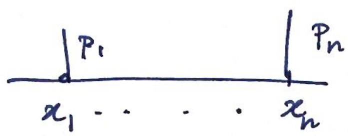
pmf of Y.

X & Y are thus two different RV's that have the same prob. mass func & same cumulative distribution fn.

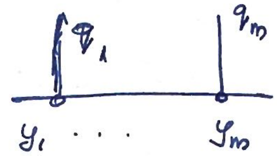
Discussion: $M_X(t) = M_Y(t) \ \forall t$ for RVs X, Y.

$$\Rightarrow E(X) = E(Y) \Rightarrow \mu_X = \mu_Y \quad \text{same means}$$
$$\Rightarrow E(X^2) = E(Y^2) \Rightarrow \sigma_X^2 = \sigma_Y^2 \quad \text{same variances.}$$

$\underline{X}$



$\underline{Y}$



$x_n > x_1$ largest value

$X = x_i \quad i = 1$ to n

$y_m > y_1$

$Y = y_i \quad i = 1$ to m

$$E(e^{tX}) = p_1 e^{tx_1} + \dots + p_n e^{tx_n} = M_X(t)$$
$$E(e^{tY}) = q_1 e^{ty_1} + \dots + q_m e^{ty_m} = M_Y(t)$$
$\Rightarrow E(e^{tX}) = E(e^{tY}) \ \forall t.$

What happens as $t \to \infty$

$$M_X(t) = E(e^{tX}) \simeq p_n e^{tx_n}$$
$$M_Y(t) = E(e^{tY}) \simeq q_m e^{ty_m}$$

For $M_X(t) = M_Y(t) \Rightarrow n = m; \ x_n = y_m \ \& \ p_n = q_m$

Subtract off the last terms and apply the argument again $\Rightarrow x_i = y_i \ \& \ p_i = q_i$

$$p_n e^{tx_n} \sim q_m e^{ty_m} \quad \text{as } t \to \infty$$

$$n = m \Rightarrow p_i = q_i \ ; \ x_i = y_i \qquad 1 \leq i \leq n$$

$\therefore$ X & Y have the same distribution

Poisson Distribution: PMF: $P(X=k) = \dfrac{\lambda^k e^{-\lambda}}{k!}$

$$M_X(t) = \sum_{n=0}^{\infty} P(X=n)e^{tn} = \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} e^{tn} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda e^t)^n}{n!}$$

$$M_X(t) = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

$$M_X'(t) = \lambda e^t e^{\lambda(e^t-1)} \qquad M_X'(0) = \mu_X = e^{-\lambda} e^{\lambda} \lambda = \lambda$$

$$M_X''(t) = \lambda e^{-\lambda} (e^t e^{\lambda e^t})' \qquad M_X''(0) = \lambda e^{-\lambda} \{ e^{\lambda} + \lambda e^{\lambda} \}$$
$$= \lambda e^{-\lambda} \{ e^t e^{\lambda e^t} + \lambda e^{\lambda e^t} e^t \} \qquad = \lambda + \lambda^2 = E[X^2]$$

$$\therefore \sigma^2 = E[X^2] - (E(X))^2 = \lambda + \lambda^2 - \lambda^2$$
$$\sigma^2 = \lambda$$

Standard Normal Distribution : $z \sim N(0,1)$, $M_z(t) = ?$

$$M_z(t) = E[e^{tz}] = \int_{-\infty}^{\infty} e^{tx} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}}_{f_z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}} dx$$

$f_z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ probability density function of z.

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

$$tx - x^2/2 = -x^2/2 + tx$$
$$= -1/2 [x^2 - 2tx].$$
$$= -1/2 [(x-t)^2 - t^2]$$

$$M_z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}[(x-t)^2 - t^2]} dx$$

$$M_z(t) = \frac{1}{\sqrt{2\pi}} e^{t^2/2} \int_{-\infty}^{\infty} e^{-1/2(x-t)^2} dx$$

If $u = x - t$ $du = dx$

$$M_z(t) = \ \ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2(x-t)^2} dx = \frac{+1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2 u^2} du = 1$$

$$\boxed{M_z(t) = e^{t^2/2}}$$

$$\because \int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$$

Normal Distribution : X with parameter μ, σ^2

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

We can take X to be $\sigma z + \mu$ where z is the standard normal.

$$M_X(t) = M_{\sigma z+\mu}(t) = E[e^{t(\sigma z+\mu)}] = E[e^{t\sigma z} \underbrace{e^{t\mu}}_{\text{constant}}]$$

$$M_X(t) = e^{t\mu} E[e^{t\sigma z}]$$

$$M_{\sigma z+\mu}(t) = e^{t\mu} E[e^{(t\sigma)z}] = e^{t\mu} M_z(\sigma t)$$

$\{ \because M_z(t) = E[e^{tz}]$
$\because M_z(\sigma t) = E[e^{t\sigma z}] \}$

$$M_{\sigma z+\mu}(t) = e^{t\mu} M_z(\sigma t)$$

$$\boxed{M_{\sigma z+\mu}(t) = e^{t\mu} e^{\sigma^2 t^2/2}}$$

Properties of MGF

① $M_{ax}(t) = E[e^{t(ax)}] = E[e^{(ta)x}] = M_x(at)$

② If X & Y are independent R.V.'s

$$M_{(X+Y)}(t) = M_X(t) \, M_Y(t).$$

If X & Y are independent random variables,

$$M_{(X+Y)}(t) = E[e^{t(X+Y)}] = E[e^{tX} \cdot e^{tY}].$$

$$= E[e^{tX}]\, E[e^{tY}] \quad \text{since } X \,\&\, Y \text{ are independent } e^{tX} \text{ and } e^{tY} \text{ are also independent RVs.}$$

$$= M_X(t)\, M_Y(t)$$

Because e^{tX} & e^{tY} are power series of RVs X & Y.

③ $$M_{X+c}(t) = E[e^{t(X+c)}] = E[e^{tX} e^{tc}]$$

$$= e^{tc} E[e^{tX}] = e^{tc} M_X(t). \qquad \{\because E[bY] = bE[Y]\}$$

Moment generating function of the Normal Distribution.

X: RV normal, with prob. density fn $f_X(x)$

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

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad M_X(t) = ?$$

Z: $N(0,1)$ standard normal has the pdf:

$$f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad M_Z(t) = e^{t^2/2}$$

$$X = \sigma Z + \mu \qquad \because Z = \frac{X-\mu}{\sigma}$$

The cumulative distribution function CDF:

$$F(x) = P(\sigma Z + \mu \leq x) = P\left(Z \leq \frac{x-\mu}{\sigma}\right)$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{(x-\mu)}{\sigma}} e^{-t^2/2}\, dt \quad \leftarrow \text{density of } Z$$

$$F(x) = P(\sigma Z + \mu \leq x) \leftarrow \text{cdf of } X$$

$$F'(x) = \text{pdf of } \sigma Z + \mu = X.$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-t^2/2}\, dt$$

$$\frac{d}{dx}\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}$$

$$\therefore F'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot \frac{1}{\sigma}$$

we want $M_x(t)$. $\because X = \mu + \sigma Z$

$\therefore M_{\sigma Z+\mu}(t) = \underset{\uparrow \text{ property ②}}{e^{t\mu} M_{\sigma Z}(t)} = \underset{\uparrow \text{ property ①}}{e^{\mu t} M_Z(\sigma t)}$ $\qquad M_Z(t) = e^{t^2/2}$

$M_X(t) = M_{\sigma Z+\mu}(t) = e^{\mu t} e^{\sigma^2 t^2/2}$

$$\boxed{M_X(t) = e^{\sigma^2 t^2/2 + \mu t}}$$

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

$M_X'(t) = e^{\sigma^2 t^2/2 + \mu t}[\mu + \sigma^2 t]$

$$\boxed{M_X'(t) = M_X(t)(\sigma^2 t + \mu)}$$

$M_X'(0) = M_X(0)(\sigma^2 \cdot 0 + \mu) = \mu M_X(0). = \mu$ (mean).

$\because M_X(0) = E[e^{0x}] = E(1) = 1$

(or) $M_X(0) = e^{\sigma^2 0^2/2 + \mu \cdot 0} = e^0 = 1$

$$\boxed{M_X''(t) = M_X'(t)(\sigma^2 t + \mu) + M_X(t) \cdot \sigma^2}$$

$M_X''(0) = \underbrace{M_X'(0)}_{\mu}[\sigma^2 0 + \mu] + \underbrace{M_X(0)}_{1} \sigma^2$

$M_X''(0) = \mu^2 + \sigma^2 = E[X^2]$

$\therefore Var[X] = M_X''(0) - (E[X])^2 = E[X^2] - (E[X])^2$

$var[X] = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$ (variance).

Theorem. $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$. Let X, Y be independent random variables

then, $X+Y$ is normal $\Rightarrow X+Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$

Note: This would be a difficult exercise to prove. the trick is to recognize that:

- Moment generating fuctions completely determines the distribution

Proof. X and Y are independent normal RVs.

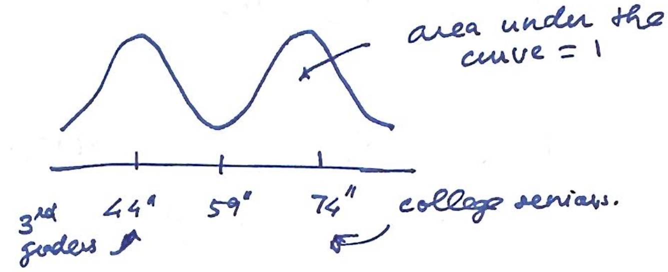
$$M_{X+Y}(t) = M_X(t)\, M_Y(t) = e^{(\sigma_X^2 t^2/2 + \mu_X t)} e^{(\sigma_Y^2 t^2/2 + \mu_Y t)}$$

$$M_{X+Y}(t) = e^{\left[\underbrace{\left(\frac{\sigma_X^2+\sigma_Y^2}{2}\right)}_{\text{variance}} t^2 + \underbrace{(\mu_X+\mu_Y)}_{\text{mean.}} t\right]}$$

This is the MGF of a RV with $\mu = \mu_X + \mu_Y$ and variance $\sigma^2 = \sigma_X^2 + \sigma_Y^2$

Example.

$\mu = 59''$ mean of the population



10000 3rd grade male basket ball players
10000 college sr. male basket ball players
} lump them together to form a population

Random variable X: height

Choose a student at random. Is X normal?
Measure his height and put him back into the population (replace him)

Why replace? Because only then would successive sampling from a population be INDEPENDENT

$\therefore$ $X \to$ height.

Start sampling: $X_1, X_2, \ldots X_n, \ldots$ are RV (independent RV).

Each of X_i is a copy of X as we are sampling from X.

$\Rightarrow \mu_{X_i} = 59$ as $\mu_X = 59$

$$P(X_5 < 59) = \frac{1}{2} = P(X_6 > 59)$$

$$\bar{X}_2 = \frac{X_1 + X_2}{2}$$ Take a sample of size 2 and take their average.

$\bar{X}$ bounces around as it is a small sample

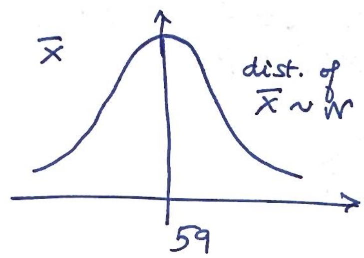
$\overline{X}_3 = \dfrac{x_1+x_2+x_3}{3}$ (still bounces around, but better than last time (with two))

$\uparrow$ sample mean of 3 values

In $n \geq 30$

$$\overline{X} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$\overline{X} = \overline{X}_n$



The values assumed by $\overline{X}$ will hover around 59, with not much bouncing.

The average smooths things out.

$\overline{X}$ is approximately normal with mean $\mu_{\overline{x}} = 59 =$ mean of the original population

Central Limit Theorem

Like, all roads leads to Rome,

All averages of a population in the limit tends to a normal distribution.