Notes

Basic Rules of Probability.

(i) For any event A, we have $0 \le P(A) \le 1$; $P(\phi)=0$: $P(\Omega)=1$

(ii) $P\left(\bigcup_k A_k\right) \le \sum_k P(A_k)$, for any countable collection of events A_k

(iii) $P\left(\bigcup_k A_k\right) = \sum_k P(A_k)$ if A_k is a countable collection of pairwise disjoint events

Inclusion-Exclusion formula.

$$P\left(\bigcup_{i=1}^{n} A_i\right) = S_1 - S_2 + S_3 - \cdots + (-1)^{n-1} S_n.$$

$$S_k = \sum_{1 \le i_1 \le i_2 \cdots < i_k \le n} P(A_{i_1} \cap A_{i_2} \cap \cdots A_{i_k})$$

Bonferroni's Inequality.

Let $A_1, A_2 \ldots A_n$ be events in a probability space (Ω, p) and let $A = A_1 \cup \cdots \cup A_n$.
Then $\boxed{S_1 - S_2 \le P(A) \le S_1}$

$$P(A) \le S_1 - S_2 + \cdots + S_m \text{ if } m \text{ is odd}$$

$$P(A) \le S_1 - S_2 + \cdots - S_m \text{ if } m \text{ is even}$$

$P(B) > 0$, $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Total probability rule - $P(B) = P(A_1) \cdot P(B|A_1) + \cdots P(A_n) \cdot P(B|A_n)$

Bayes rule $P(A_k|B) = \dfrac{P(A_k) \cdot P(B|A_k)}{P(A_1) \cdot P(B|A_1) + \cdots + P(A_n) P(B|A_n)}$

→ Studentized theorem

(7) Discrete Probability Distribution.

Name	PMF	CDF($F_X(t)$)	MGF($M_X(t)$)	Mean ($E(X)$)	Variance ($\sigma^2 = E[(X-\mu)^2]$)
(1) Bernoulli $Ber(p)$ $\to \{0,1\}$	$f_X(t) = \begin{cases} p & t=1 \\ 1-p & t=0 \end{cases}$	$F_X(t) = \begin{cases} 1 & t \geq 1 \\ 1-p & t \in [0,1) \\ 0 & t < 0 \end{cases}$	$pe^t - p + 1$	p	$(1-p)p$
(ii) Binomial $Bin(n,p)$ $\to [0,1]^n$	$f_X(k) = \begin{cases} {}^nC_k p^k (1-p)^{n-k} & 0 \leq k \leq n \\ 0 & o/w \end{cases}$	$F_X(t) = \begin{cases} \sum_{r=0}^{k} {}^nC_r p^r (1-p)^{n-r} & k \leq t < k+1 \\ 0 & o/w \end{cases}$	$(p(e^t - 1) + 1)^n$	np	$n(1-p)p$
(iii) Geometric $Geo(p)$ $\to [0,1]^n$	$f_X(t) = \begin{cases} p(1-p)^t & t \geq 0 \\ 0 & o/w \end{cases}$	$F_X(t) = \begin{cases} 0 & \text{if } t < 1 \\ 1-(1-p)^k & \text{if } k \leq t < k+1 \\ 1 & t \geq n. \end{cases}$	$\dfrac{p}{1-(1-p)e^t}$	$\frac{1}{p} - 1$	$\dfrac{1-p}{p^2}$
(iv) Poisson $Pois(\lambda)$ $[0,1]^n$	$f_X(k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$	$F_X(k) = \sum_{r=0}^{k} e^{-\lambda} \dfrac{\lambda^r}{r!}$	$e^{\lambda(e^t - 1)}$	λ	λ
(v) Hypergeometric ~~Hypergeo(n,m)~~ Hypergeo (n,m,N)	$f_X(k) = \dfrac{{}^mC_x \, {}^{N-m}C_{n-x}}{{}^NC_n}$ $0 \leq x \leq m$, $m+n-N \leq x \leq m$, $0 \leq x \leq n$, $m+n-N \leq x \leq n$	$F_X(k) = \sum_{x=0}^{k} \dfrac{{}^mC_x \, {}^{N-m}C_{n-x}}{{}^NC_n}$	-	$\dfrac{mn}{N}$	$\dfrac{\frac{mn(N-m)(N-n)}{N-1}}{N} = \dfrac{mN(1-\frac{m}{N})(N-n)}{(N-1)N}$
Negative Binomial (n,p)	$P(X=x) = p^n (1-p)^x \, {}^{n+x-1}C_{n-1}$, $x \geq 0$	$\sum P$	$\left(\dfrac{p}{1-(1-p)e^t}\right)^n$	$\dfrac{n(1-p)}{p}$	$\dfrac{n(1-p)}{p^2}$

(8) $X_n \xrightarrow{P} a$: $P(|X_n - a| > \epsilon) = 0$

$X_n \xrightarrow{D} Y$: $\lim_{n\to\infty} F_{X_n} = F_Y$, $\lim_{n\to\infty} M_{X_n} = M_Y$

Slutsky: $X_n \xrightarrow{D} X$, $Y_n \xrightarrow{P} C$, $X_n Y_n \xrightarrow{D} XC$

WLLN: $\bar{X}_n \to \mu$ (No Chebyshev needed)

Laplace: $\frac{1}{2} e^{-|x|}$

Name	PDF $f_X(t)$	CDF $F_X(t)$	Moment generating fn $M_X(t)$	Mean $E[X]$	Variance $E[(X-\mu)^2]$
(i) Uniform Unif[a,b]	$f(t)=\begin{cases}\frac{1}{b-a} & t\in(a,b)\\ 0 & o/w\end{cases}$	$F_X(t)=\begin{cases}0 & t\le a\\ \frac{t-a}{b-a} & t\in(a,b)\\ 1 & t\ge b\end{cases}$	$\frac{e^{bt}-e^{at}}{t(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
(ii) Exponential ~ Exp(λ)	$f(t)=\begin{cases}\lambda e^{-\lambda t} & t>0\\ 0 & o/w\end{cases}$	$F(t)=\begin{cases}0 & t\le 0\\ 1-e^{-\lambda t} & t>0\end{cases}$	$\frac{\lambda}{\lambda-t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
(iii) Normal ~ $N(\mu,\sigma^2)$	$\varphi_{\mu,\sigma^2}(t)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	$\Phi_{\mu,\sigma^2}(t)=\int_{-\infty}^{t}\varphi_{\mu,\sigma^2}(u)\,du$	$e^{[(\sigma^2t^2)/2+t\mu]}$	μ	σ^2
(iv) Gamma ~ Gamma(ν,λ)	$f(t)=\begin{cases}\frac{1}{\Gamma(\nu)}\lambda^{\nu}t^{\nu-1}e^{-\lambda t} & t>0\\ 0 & o/w\end{cases}$ $\Gamma(\nu)=\int_0^{\infty}t^{\nu-1}e^{-t}dt$	$F(t)=\begin{cases}0 & t\le 0\\ \int_0^t f(u)\,du & t>0\end{cases}$	$(1-t/\lambda)^{-\nu}$	$\frac{\nu}{\lambda}$	$\frac{\nu}{\lambda^2}$
(v) Beta ~ Beta(α,β)	$f(t)=\begin{cases}\frac{1}{B(\alpha,\beta)}t^{\alpha-1}(1-t)^{\beta-1} & t\in(0,1)\\ 0 & o/w\end{cases}$ $B(\alpha,\beta)=\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}$	$F(t)=\begin{cases}0 & t\le 0\\ \int_0^t f(u)\,du & t\in(0,1)\\ 0 & t\ge 1\end{cases}$	-	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
(vi) Cauchy(a,λ)	$f(t)=\frac{\lambda}{\pi(\lambda^2+(t-a)^2)}$	$F(t)=\frac{1}{2}+\frac{1}{\pi}\tan^{-1}\left(\frac{t-a}{\lambda}\right)$	undefined	-	-

(2) Properties

(i) Gamma fn: $\Gamma(\nu+1)=\nu\Gamma(\nu)$

$$\Gamma(1/2)=\int_0^{\infty}x^{-1/2}e^{-x}dx=\sqrt{2}\int_0^{\infty}e^{-y^2/2}dy \; ; \; \Gamma(1/2)=\sqrt{\pi}$$

(ii) Beta fn: $B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

(iii) $N(0,1)$ → Standard Normal dist.

$$\boxed{\Phi_{\mu,\sigma^2}(t)=\Phi\left(\frac{t-\mu}{\sigma}\right)} \qquad \varphi(t)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2}$$

① Markov's Inequality -

X be a non-negative RV with finite expectation.

$$\boxed{P\{X \geq t\} \leq \frac{1}{t} E[X]}$$

② X - be a non-ve RV with finite $E[X^p]$

$$\boxed{P\{X \geq t\} \leq \frac{1}{t^p} E[X^p]}$$

③ Chebyshev's Inequality.

X be a RV with finite second moment, $\mu = E[X]$

$$\boxed{P[|X-\mu| \geq t] \leq \frac{1}{t^2} \operatorname{var}(X)}$$

④ Chernoff's Inequality.

X be a RV with finite exponential moments.

$$\boxed{P\{X \geq t\} \leq e^{-\lambda t} E[e^{\lambda X}]} \quad \text{for } \lambda > 0$$

⑤ Jensen's Inequality.

Let $I \subseteq \mathbb{R}$ be an interval and let $\varphi : I \to \mathbb{R}$ be a twice differential f^n st its φ'' is continuous on I and $\varphi''(x) \geq 0$ $\forall x \in \mathbb{R}$ (φ is convex). Let X be a RV with $S_X \subseteq I$ and finite expectation.

$$\boxed{E[\varphi(X)] \geq \varphi(E[X])}$$

⑥ AM - GM - HM Inequality.

Let X be a RV with $S_X \subseteq (0, \infty)$. Then $E[X]$ is called AM of X

$e^{E(\ln X)}$ is called GM of X

$\frac{1}{E(1/X)}$ is called HM of X

$$\boxed{E[X] \geq e^{E[\ln X]} \geq \frac{1}{E[1/X]}}$$

Notes.

CDF (Cumulative Distribution function)

Def: The Cumulative distribution f^n of a R.V X is a f^n $F_X : \mathbb{R} \to [0,\infty)$ defined by $F_X(x) = P(X \le x)$

The CDF has the following properties:

① F_X is non decreasing and hence has only jump discontinuities.

② $\lim_{x \to \infty} F_X(x) = 1 \quad \lim_{x \to -\infty} F_X(x) = 0$

③ $\lim_{h \to 0} F_X(x+h) = F_X(x) \;\; \forall\, x \in \mathbb{R}$

④ $\lim_{h \to 0}$ $P(X) = F_X(x) - F_X(x^-) \;\; \forall\, x \in \mathbb{R}$

⑤ $P(a < X \le b) = F_X(b) - F_X(a)$

$P(a \le X \le b) = F_X(b) - F_X(a^-)$

$P(a < X < b) = F_X(b^-) - F_X(a)$

$P(a \le X < b) = F_X(b^-) - F_X(a^-)$

$$\int_{t=a}^{t=b} f(x(t), y(t)) \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

Line integral

Change of variable

① pmf: $\quad h : \mathbb{R} \to \mathbb{R}$

$f_Z(x) = \begin{cases} f_X(h^{-1}(z)) & \text{if } z \in S_Z \\ 0 & o/w \end{cases}$

$Z = h(X)$

② pdf

- Identify S_X
- → Look at h (split S_X into parts where h is strictly ↑ing)

$\therefore f_Z(t) = \sum_{j=1}^{k} f_X(h_j^{-1}(t)) \left|\frac{d}{dt} h_j^{-1}(t)\right| I_{h(S_j)}(t)$

Probability mass function (pmf)

A RV is said to have discrete distribution if ∃ an atmost countable set $S_X \subset \mathbb{R}$ st $P(X=x) > 0$ $\forall\, x \in S_X$ and $\sum_{x \in S_X} P(X=x) = 1$. S_X is called the support of X.

pmf $=$ $f_X(x) = \begin{cases} P(X=x) & \text{if } x \in S_X \\ 0 & o/w \end{cases}$

CDF $=$ • $F_X(x) = \sum_{y \le x} f_X(y)$

• $f_X(x) = F_X(x) - F_X(x^-)$

Properties of pmf:

① $f_X(x) \ge 0 \;\; \forall\, x \in \mathbb{R}$

② $\sum_{x \in S_X} f_X(x) = 1$

and S_X = support of Random variable.

$S_X = \{x : f_X(x) > 0\}$ is at most countable (countable or finite)

Theorem:

Suppose a real valued f^n $h : \mathbb{R} \to \mathbb{R}$ satisfies the following:

① $h(x) \ge 0 \;\; \forall\, x \in \mathbb{R}$ $\quad D = \{x : h(x) > 0\}$ is atmost countable

② $\sum_{x \in D} h(x) = 1$

$\therefore h(\cdot)$ is a pmf.

Examples (pmf)

① (Bernouli distribution) $X \sim Bernoulli(p)$: $S_X = \{0,1\}$
$f_X(0) = 1-p$; $f_X(1) = p$

② (Binomial distribution) $X \sim Bin(n,p)$: $S_X = \{0,1,\dots n\}$
$f_X(K) = {}^nC_K p^K (1-p)^{n-K}$

③ (Geometric distribution) $X \sim Geo(p)$: $S_X = \{0,1\dots\}$
$f_X(K) = p(1-p)^K$

④ (Poison Distribution) $X \sim Poi(\lambda)$ $(\lambda > 0)$: $S_X = \{0,1\dots\}$
$f_X(K) = \dfrac{e^{-\lambda}\lambda^K}{K!}$

⑤ (Hypergeometric Distribution) $X \sim Hypergeo(b, w, m)$: S_X $\{\max(0, m-w), \dots, b\}$.
$f_X(K) = \dfrac{{}^bC_K \, {}^wC_{m-K}}{{}^{b+w}C_m}$

④ Probability Density function (pdf)

A RV is said to have a Continuous distribution if $\exists$ a non-ve integrable f^n $f_X : \mathbb{R} \to [0,\infty)$ st ~~CRV~~: $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$ $\forall x \in \mathbb{R}$

the set $S_X = \{x \in \mathbb{R} : f_X(x) > 0\}$ is called support of X.

Here X is CRV (continuous Random Variable)

- For CRV X $P(X=a) = 0$ $\forall a \in \mathbb{R}$
- CDF of PDF is cont^n.
- PDF is not unique. (At discontinuous pt, can take any value)
- S_X isn't unique.
- $P(a \le X \le b) = \int_a^b f_X(t)\,dt$
- $f_X(x)$ is not $P(X=x)$

Properties of PDF:

① $f_X(x) \ge 0$ $\forall x \in \mathbb{R}$ ⑪ $\int_{-\infty}^{\infty} f_X(x) = 1$

Thm: Suppose a real valued f^n $g: \mathbb{R} \to \mathbb{R}$ satisfies the following cond^n:

① $g(x) \ge 0$ $\forall x \in \mathbb{R}$
⑪ $\int_{-\infty}^{\infty} g(x)\,dx = 1$

$\Big\}$ g is pdf of some cont^n random variable

Example – ① (Exponential distr.): $Exp(\lambda)$ $F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x > 0 \\ 0 & o/w \end{cases}$

⑪ (Uniform distr.) $U(a,b)$ $F_X(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \le x < b \\ 1 & \text{if } x \ge b. \end{cases}$

ⅲ (Normal distr.) $N(\mu, \sigma^2)$

~~$F_X(x) = \int_{-\infty}^{x}$ 0 if $x < a$~~

$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt$ if $-\infty < x < \infty$.

(*) Change of Variable.

① CDF technique -- Distribution $Z = h(X)$ can be determined by computing the dist func.

$$F_Z(z) = P(Z \le z) = P(h(X) \le z)$$

$$\Rightarrow F(X \le h^{-1}(z))$$

② For pmfs :- Find S_X, S_Z

- For each element in S_Z, find elements in S_X and sum their probabilities.

$$f_Z(z) = \begin{cases} \sum_{x \in A_z} f_X(x) & \text{if } z \in S_Z \\ 0 & \text{o/w} \end{cases}$$

③ for pdfs - find S_X

$Z = h(X)$

- divide S_X into interval s.t $h(x)$ is monotonic there,

- $$f_Z(z) = \sum_{j=1}^{k} f_X(h_j^{-1}(z)) \left| \frac{d}{dt} h_j^{-1}(z) \right| I_{h(S_j)}(t)$$

(*) Mean - $E[X] = \begin{cases} \sum_x x f(t) & \text{if } \sum_x |x| f(x) < \infty \text{ for } f \text{ being pmf.} \\ \int_{-\infty}^{\infty} x f(x) & \text{if } \int_{-\infty}^{\infty} |x| f(x) < \infty \text{ for } f \text{ being pdf} \end{cases}$

Properties - ① $E[aX + bY] = aE[X] + bE[Y]$

② If $X \ge Y$ ($X(w) \ge Y(w)$ for all w) $\Rightarrow E[X] \ge E[Y]$

③ If $\varphi : \mathbb{R} \to \mathbb{R}$

$$E[\varphi(X)] = \begin{cases} \sum_t \varphi(t) f(t) & \text{if } f \text{ is pmf and } \sum_t |\varphi(t)| f < \infty \\ \int_{-\infty}^{\infty} \varphi(t) f(t) & \text{if } f \text{ is pdf and } \int_{-\infty}^{\infty} |\varphi| f < \infty \end{cases}$$

(*) Moments : $\mu_k' = E[X^k]$

(*) Central Moments $\Rightarrow \mu_k = E[(X-\mu)^k]$; where $\mu = E[X]$

(*) Variance $\Rightarrow \sigma^2 = \mu_2 = E[(X-\mu)^2] = E[X^2] - E[X]^2$

(*) $s.d(X) = \sqrt{Var(X)}$

(*) Median: $F_X(m^-) \le \frac{1}{2} \le F_X(m)$ ie.

(*) Coefficient of variation $(C.V(X)) = \dfrac{sd(X)}{|E(X)|}$

(*) Mean absolute deviation (m.a.d) $= E[|X - \text{median}(X)|]$

NEXT 6.2%

(*) Mode :- Value that occurs at highest pmf/pdf ; $f_X(m_0) = \sup\{f_X(x) : x \in S_X\}$

(*) Quartile deviation

$$F_X(q_1^-) \le \frac{1}{4} \le F_X(q_1)$$

$$F_X(q_3^-) \le \frac{3}{4} \le F_X(q_3)$$

Deviation: $q_3 - q_1$

(*) Coefficient of quartile deviation $= \frac{q_3 - q_1}{q_3 + q_1}$

(*) Standardization:

Standardised Variable: $Z = \frac{(X-\mu)}{\sigma}$; $\mu = E[X]$

$\sigma = \sqrt{E[X^2] - E[X]^2}$

(*) Skewness.

$$\beta_1 = E(Z^3) = \frac{E[(X-\mu)^3]}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}}$$

Yule coefficient of skewness: $\frac{q_3 - 2m + q_1}{q_3 - q_1}$ (β_2)

(*) Kurtosis: $\gamma_1 = E(Z^4) = \frac{E[(X-\mu)^4]}{\sigma^4} = \frac{\mu_4}{\mu_2^2}$

Excess kurtosis $(\gamma_2) = \gamma_1 - 3$

(*) MGF

Define $M_X : A \to \mathbb{R}$ by $M_X(t) = E[e^{tX}]$ $t \in A$

(mgf exists if $\exists$ a st $(-a, a) \subseteq A$ or if mgf is finite in interval containing 0)

If mgf exists - ① $\mu_r' = E(X^r)$ is finite for all $r \in \mathbb{N}$

② $\mu_r' = E[X^r] = \frac{d^r}{dt^r} M_X(t) \Big|_{t=0}$

③ $M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu_r'$ with $t \in (-a, a)$

$$\text{MGF} \xrightarrow[\frac{d}{dt}\big|_{t=0}]{} E[X] \xrightarrow[\frac{d^2}{dt^2}\big|_{t=0}]{} E[X^2]$$

(*) $\psi_X(t) = \ln M_X(t)$ $\psi_X(t) : (-a, a) \to \mathbb{R}$.

$t \in (-a, a)$

$\mu_1' = E(X) = \psi_X^{(1)}(0)$

$Var(X) = \psi_X^{(2)}(0)$

Moment generating fn. $M_X(t) = E\left(e^{\sum t_i x_i}\right)$

- $M_X(0_p) = 1$
- $E(x_1^{k_1} x_2^{k_2} \cdots x_p^{k_p}) = \left[\dfrac{\partial^{k_1+k_2+\cdots+k_p}}{\partial t_1^{k_1} \partial t_2^{k_2} \cdots \partial t_p^{k_p}} M_X(t)\right]_{t=0_p}$

$Cov(x_i, x_j) = E((x_i - E(x_i))(x_j - E(x_j)))$
$= E(x_i x_j) - E(x_i)E(x_j)$

- $Cov(x_i, x_j) = \dfrac{\partial^2}{\partial t_i \partial t_j} \varphi_X(t)\Big|_{0_p}$, $\varphi_X(t) = \ln M_X(t)$

Correlation $= \dfrac{Cov(X,Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$

$Cor(X,Y) > 0 \rightarrow$ ↑es in $X \Rightarrow$ ↑es in Y ($Cor(X,Y) = 1 \Rightarrow X = Y$)
$Cor(X,Y) < 0 \rightarrow$ ↑ in $X \Rightarrow$ ↓ in Y ($Cor(X,Y) = -1 \Rightarrow X = -Y$)
$Cor(X,Y) = 0 \rightarrow$ (X, Y ~~are~~ may or may not be independent)

Property:
- $Cov(aX_1 + bX_2, Y) = a\,Cov(X_1, Y) + b\,Cov(X_2, Y)$
  $Cov(X, aY_1 + bY_2) = a\,Cov(X, Y_1) + b\,Cov(X, Y_2)$
- Symmetry — $Cov(X,Y) = Cov(Y,X)$
- Positivity $Var(X) \geq 0$

Cauchy Schwarz inequality . $Cov(X,Y)^2 \leq Var(X) \cdot Var(Y)$

Bivariate Normal Distribution.

$$f(x,y) = \frac{\sqrt{ab-c^2}}{2\pi} e^{-\frac{1}{2}\left[a(x-u)^2 + b(y-v)^2 + 2c(x-u)(y-v)\right]}$$

let $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \Rightarrow f(x,y) = \dfrac{1}{2\pi\sqrt{\det(\Sigma)}} e^{-\frac{1}{2} t \Sigma^{-1} t}$

$$\rightarrow \boxed{M_{xy}(t_1, t_2) = e^{t_1\mu_1 + t_2\mu_2 + \frac{1}{2}(t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2t_1t_2\sigma_1\sigma_2\rho)}}$$

$\mu_1 \rightarrow E(X)$; $\mu_2 \rightarrow E(Y)$, $\sigma_1^2 \rightarrow Var(X)$, $\sigma_2^2 \rightarrow Var(Y)$, $\rho = Cor(X,Y)$, $\rho\sigma_1\sigma_2 = Cov(X,Y)$

If X, Y independent $\Rightarrow Cov(X,Y) = 0$
Vice versa is true for (Multivariate normal dist)

Multivariate Normal distribution.

$Z \rightarrow$ iid ? $Z \sim N_n(0_n, I_n) \Rightarrow f_Z(z) = \left(\frac{1}{2\pi}\right)^{n/2} \exp\left(-\frac{1}{2} t' t\right)$

$\Sigma = Cov = Var$ $\quad M_Z(t) = \exp\left(\frac{1}{2} t' t\right)$

$Y \sim \Sigma^{1/2} Z + \mu$. $M_X(t) = \exp\left(t'\mu + \frac{1}{2} t' \Sigma t\right)$; $X \sim N_n(\mu, \Sigma)$

$Y \sim AX + B \Rightarrow Y \sim N_n(A\mu + B, A\Sigma A')$

$Cov(AZ) = A\,Cov(Z)\,A^T$

③ Checking Joint DF

① $\lim_{x \to -\infty} f(x,y) = \lim_{y \to -\infty} f(x,y) = 0$

② $\lim_{x \to \infty, y \to \infty} f(x,y) = 1$

③ for each $y \in \mathbb{R}$, $f(x,y)$ is right continuous
for each $x \in \mathbb{R}$, $f(x,y)$ is right continuous

④ $P(a_1 < X_1 \le a_2, b_1 < X_2 \le b_2) = F(a_1 a_2) + F(b_1 b_2) - F(a_1 b_2) - F(a_2 b_1)$

---

④ Conditional pmfs/pdfs.

$P(X_{K+1} = s_1, \ldots X_{K+l} = s_l) > 0 = \frac{f(t_1, \ldots t_K, s_1 \ldots s_l)}{g(s_1 \ldots s_l)}$ where $s_1 = K+1 \ldots s_l = K+l$

$h_{s_1 \ldots s_l}(t_1 \ldots t_K) = \frac{f(t_1 \ldots t_K, s_1 \ldots s_l)}{g(s_1 \ldots s_l)}$

④ For independent events $f_1(x|y) = f(x)$

→ If X and Y are independent ⇒ $h_x(y) = h(y)$ for all y.
→ If $h_x(y) = h(y)$ for all x, y ⇒ X, Y are independent
→ $E[H_1(X) H_2(Y)] = E(H_1(X)) \cdot E(H_2(Y))$ for any fn's $H_1, H_2 : \mathbb{R} \to \mathbb{R}$
→ If $X_1, X_2 \ldots X_p$ are independent ⇒ $M_X(t) : E\left(e^{\sum_{i=1}^{p} t_i x_i}\right) = \prod_{i=1}^{p} M_{X_i}(t_i)$
→ If X, Y are independent $Cov(X,Y) = 0$ (not vice versa)

⑱ Moments ① Joint raw moment of order $k_1 + k_2 + \ldots + k_p =$
$\varphi(x) = x_1^{k_1} x_2^{k_2} \ldots x_p^{k_p}$

$$\boxed{\mu'_{k_1 k_2 \ldots k_p} = E(x_1^{k_1} x_2^{k_2} \ldots x_p^{k_p})}$$

② Joint central moment of order $k_1 + k_2 \ldots + k_p$

$$\boxed{\mu_{k_1 k_2 \ldots k_p} = E[(X_1 - E(X_1))^{k_1} \ldots (X_p - E(X_p))^{k_p}]}$$

④ Covariance → $Cov(x_i, x_j) = E(x_i x_j) - E(x_i) E(x_j)$

④ Properties - • $E\left(\sum_{1}^{p_1} a_i x_i\right) = \sum_{1}^{p_1} a_i E(x_i)$

• $Cov\left(\sum_{1}^{p_1} a_i x_i, \sum_{1}^{p_2} b_j Y_j\right) = \sum_{1}^{p_1} \sum_{j}^{p_2} a_i b_j Cov(X_i, Y_j)$

• $Var\left(\sum_{1}^{p_1} a_i x_i\right) = \sum_{1}^{p_1} a_i^2 Var(x_i) + 2 \sum\sum_{1 \le i < j \le p} Cov(X_i X_j)$

Student's theorem.

① Let $X_1, X_2 \cdots X_n \xrightarrow{iid} N(\mu, \sigma^2)$

$\Rightarrow$ ① $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$, where $\bar{X}_n = \frac{1}{n} \sum_{1}^{n} X_K$

② $\frac{n W_n}{\sigma^2} \sim \chi_n^2$, where $W_n = \frac{1}{n} \sum_{1}^{n} (X_K - \mu)^2$

③ $\frac{(n-1) S_n^2}{\sigma^2} \sim \chi^2_{n-1}$ where $S_n^2 = \frac{1}{(n-1)} \sum_{1}^{n} (X_K - \bar{X}_n)^2$

④ $\bar{X}_n$ and S_n^2 are independent.

⑤ $\frac{\bar{x} - \mu}{S_n / \sqrt{n}} = \frac{\sqrt{n}(\bar{x} - \mu)}{S_n} \sim t_{n-1}$

$\rightarrow Cov(X_i, X_j) = E((X_i - E(X_i))(X_j - E(X_j))) = E(X_i X_j) - E(X_i) E(X_j)$

$\rightarrow E[\sum a_i X_i] = \sum a_i E[X_i]$

⊛ $Cov(X_1, X_2) = 0$ ($X_1, X_2 \to$ independent), vice-versa not true

$\Rightarrow$ true for Normal

$$Cov(X_i, X_j) = E((X_i - E(X_i))(X_j - E(X_j)))$$
$$= E(X_i X_j) - E(X_i)E(X_j)$$

⊛ We can find $E[X]$ from joint pdfs as well

⊛ For indepent joint pdfs, try splitting into individual fns

⊛ $M_X(0, 0, t_i, 0 \dots t_j) = E(e^{t_i X_i + t_j X_j})$

⊛ $Cov(X, X) = Var(X)$

⊛ Cauchy Schwartz: $Cov(X, Y)^2 \leq Var(X) \cdot Var(Y)$

Equality if $Y = aX + b$

⑤ X_i = ind.

$$M_X(t) = \prod M_{X_i}(t)$$

$X_n \to X$, $Y_n \to c$ (const)

⊛ WLLN $\left[\bar{X} \to \mu, \frac{\sum X_i^2}{n} \to E[X^2]\right]$, Slutsky $\left(X_n + Y_n \xrightarrow{D} X + c,\ X_n \cdot Y_n \xrightarrow{D} Xc\right)$, CMT (Continuous Mapping theorem)

$\bar{X}_n \to X_0$

$g(\bar{X}_n) \to g(X)$, g is contth

Things learnt

(1) Solving problem of min $\Rightarrow P(\min(x_1, x_2 \ldots) \geq t) = P(x_i \geq t \ \forall \ i = 1, N)$

(2) $X_n \xrightarrow{d} X \Rightarrow MGF(X_n) = MGF(X)$
and vice-versa (often makes things simpler)

(3) Given n = fixed value (large) and no pdf, then try using Central limit theorem, $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0,1) \Rightarrow \sqrt{n}(\bar{X}_n - \mu) \sim N(0, \sigma^2)$

Property of Normal

(4) Use log in case of product of RVs

Also $X \sim U(0,1)$

$Y = -\log(x_i)$

$Y \sim \exp(1)$

applied only if moments are finite

(5) For weak law of large numbers, we don't need Chebyshev's inequality.
Its in general that $E[\bar{X}_n] \to \mu$ for $n \to \infty$

where as for $\bar{X} \xrightarrow{P} \mu$
we need Chebyshev's inequality

$P(|X - \mu| > \epsilon) \leq \frac{1}{\epsilon^2} Var(X)$

(6) Average of Cauchy sequence turns out to be Cauchy.

$\cancel{X_{n+1} = \frac{n}{n+1} \bar{X}_n + \frac{1}{n+1} X_{n+1}}$ $\Rightarrow$ $X_n \sim$ Cauchy

$Y = \alpha X_1 + (1-\alpha) X_2$ $\qquad$ $X_{20} \sim$ Cauchy

$X_1 \to$ Cauchy

---

(7) Rules for Central limit theorem

(i) $t_n \to t$, $s_n \to s$ (const)
$t_n \times s_n \to ts$.

(ii) $t_n \to t \Rightarrow g(t_n) \to g(t)$ (g is cont^n)

(iii) delta method: $\sqrt{n}(Y_n - \mu) \sim N(0,1)$
$\sqrt{n}(g(Y_n) - g(\mu)) \sim g'(\mu) N(0,1)$
$\sim N(0, g'(\mu)^2)$

(8) Estimating discrete to continuous, use continuity correction $\Rightarrow P(a \leq X \leq b) = P(a - 0.5 \leq X \leq b + 0.5)$

Distribution Table.

	Normal CDF at Z			χ^2- distribution			t-distribution quantities.		
	Z	+0.00	+0.05	n=df	$\chi_n^2(0.10)$	$\chi_n^2(0.90)$	n=df	$t_n(0.10)$	$t_n(0.05)$	$t_n(0.025)$
	0	0.5	0.51994	1	2.7055	0.015791	1	3.0777	6.3138	12.706
	0.1	0.53983	0.55962	2	4.6052	0.21072	2	1.8856	2.92	4.3027
	0.2	0.57926	0.59871	3	6.2514	0.58437	3	1.6377	2.3534	3.1824 ~~4.3027~~
	0.3	0.61791	0.63683	4	7.7794	1.0636	4	1.5332	2.1318	2.7764
	0.4	0.65542	0.67364	5	9.2364	1.6103	5	1.4759	2.015	2.5706
	0.5	0.69146	0.70884	6	10.645	2.2041	6	1.4398	1.9432	2.4469
7	0.6	0.72575	0.74215	7	12.017	2.8331	7	1.4149	1.8946	2.3646
8	0.7	0.75804	0.77337	8	13.362	3.4895	8	1.3968	1.8595	2.306
	0.8	0.78814	0.80234	9	14.684	4.1682	9	1.383	1.8331	2.2622
	0.9	0.81594	0.82894	10	15.987	4.8652	10	1.3722	1.8125	2.2281
	1	0.84134	0.85814	11	17.275	5.5778	11	1.3634	1.7959	2.201
	1.1	0.86433	0.87493	12	18.549	6.3038	12	1.3562	1.7823	2.1788
	1.2	0.88493	0.89435	13	19.812	7.0415	13	1.3502	1.7709	2.1604
	1.3	0.9032	0.91149	14	21.064	7.7895	14	1.345	1.7613	2.1448
	1.4	0.91924	0.92647	15	22.307	8.5468	15	1.3406	1.7531	2.1314
	1.5	0.93319	0.93943	16	23.542	9.3122	16	1.3368	1.7459	2.1199
	1.6	0.9452	0.95053	17	24.769	10.085	17	1.3334	1.7396	2.1098
	1.7	0.95543	0.95994	18	25.989	10.865	18	1.3304	1.7341	2.1009
19	1.8	0.96407	0.96784	19	27.204	11.651	19	1.3277	1.7291	2.093
20	1.9	0.97128	0.97441	20	28.412	12.443	20	1.3253	1.7247	2.086
21	2.0	0.97725	0.97982	21	29.615	13.24	21	1.3232	1.7207	2.0796
22	2.1	0.98214	0.98422	22	30.813	14.041	22	1.3212	1.7171	2.0739
23	2.2	0.9861	0.98778	23	32.007	14.848	23	1.3195	1.7139	2.0687
24	2.3	0.98928	0.99061	24	33.196	15.659	24	1.3178	1.7109	2.0639
25	2.4	0.9918	0.99286	25	34.382	16.473	25	1.3163	1.7081	2.0595
26	2.5	0.99379	0.99461	26	35.563	17.292	26	1.315	1.7056	2.0555
27	2.6	0.99534	0.99598	27	36.741	18.114	27	1.3137	1.7033	2.0518
28	2.7	0.99653	0.99702	28	37.916	18.939	28	1.3125	1.7011	2.0484
29	2.8	0.99744	0.99781	29	39.087	19.768	29	1.3114	1.6991	2.0452
30	2.9	0.99813	0.99841	30	40.256	20.599	30	1.3104	1.6973	2.0423

(4) Bivariate Normal dist.: $f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-Q/2}$ $\quad -\infty<x<\infty, -\infty<y<\infty$; $Q = \frac{1}{(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]$

$-\infty<\mu_i<\infty, \sigma_i>0, \rho^2<1$

(5) Joint pdf of ith & jth order statistic

$$g_{ij}(y_i,y_j) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!}\left[F(y_i)\right]^{i-1}\left[F(y_j)-F(y_i)\right]^{j-i-1}\left[1-F(y_j)\right]^{n-j} f(y_i)f(y_j) & a<y_i<y_j<b \\ 0 & o/w \end{cases}$$

CDFs (Discrete)

Name	pmf	CDF

① Ber(p)

$f_X(t) = \begin{cases} p & t=1 \\ 1-p & t=0 \end{cases}$

$F = \begin{cases} 1 & t \ge 1 \\ q & t \in [0,1) \\ 0 & t<0 \end{cases}$

$\mu = p, \ \sigma^2 = (1-p)p.$

$M(t) = pe^t - p + 1$

② Bin(n,p)

$f = \begin{cases} {}^nC_k \, p^k (1-p)^{n-k} & 0 \le k \le n \\ 0 & o/w \end{cases}$

$F = \begin{cases} \sum_{r=0}^{k} {}^nC_r \, p^r (1-p)^{n-r} & 0 \le k \le n \\ 0 & o/w \end{cases}$

$\mu = np, \ \sigma^2 = n(1-p)p.$

$M(t) = (p(e^t-1)+1)^n.$

③ Geo(p)

$f_X(t) = \begin{cases} p(1-p)^t & t \ge 0 \\ 0 & o/w \end{cases}$

$F_X(t) = \begin{cases} 0 & t<1 \\ 1-(1-p)^k & k \le t < k+1 \\ 1 & t \ge n \end{cases}$

$\mu = \frac{1}{p} - 1, \ \sigma^2 = \frac{1-p}{p^2}$

$M(t) = \frac{p}{1-(1-p)e^t}$

④ Pois(λ)

$f_X = \frac{e^{-\lambda}\lambda^k}{k!}$

$F = \sum f_X.$

$\mu = \lambda, \ \sigma^2 = \lambda.$

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

⑤ Neg.Bin(n,p)

$f = p^n (1-p)^x \, {}^{n+x-1}C_{n-1} \quad x \ge 0$

$F = \sum F$

$\mu = \frac{n(1-p)}{p}$

$\sigma^2 = \frac{n(1-p)}{p^2}$

$M(t) = \left(\frac{p}{1-(1-p)e^t}\right)^n$

⑤ Hypergeo(n,m,N)

$f_X = \frac{{}^mC_x \, {}^{N-m}C_{n-x}}{{}^NC_n}$, $0 \le x \le m$; $m+n-N \le x \le m$; $0 \le x \le n$; $m+n-N \le x \le n$

$F = \sum f$

$\mu = \frac{mn}{N}, \ \sigma^2 = \frac{mN(1-\frac{m}{N})(N-n)}{(N-1)N}$

① Uniform(α,β)

$f = \frac{1}{b-a} \quad a<x<b.$

$F = \begin{cases} 0 & t \le a \\ \frac{t-a}{b-a} & (a,b) \\ 1 & t \ge b \end{cases}$

$\mu = \frac{a+b}{2}, \ \sigma^2 = \frac{(b-a)^2}{12}$

$M(t) = \frac{e^{bt} - e^{at}}{(b-a)t}, \ -\infty<t<\infty$

② Exponential

$Exp(\lambda) \Rightarrow \Gamma(1, 1/\lambda)$

$f = \begin{cases} \lambda e^{-\lambda x} & x>0 \\ 0 & o/w \end{cases}$

$F = \begin{cases} 0 & t \le 0 \\ 1-e^{-\lambda t}, & t>0 \end{cases}$

$\mu = \frac{1}{\lambda}, \ \sigma^2 = \frac{1}{\lambda^2}$

$M(t) = \frac{\lambda}{\lambda - t}$

③ Normal → N(μ,σ²)

$\phi_{\mu,\sigma^2}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$

$\Phi_{\mu,\sigma^2}(t) = \int_{-\infty}^{t} \phi$

$\mu = \mu, \ \sigma^2 = \sigma^2$

$M(t) = e^{\left[\frac{\sigma^2 t^2}{2} + t\mu\right]}$

④ Gamma Γ(α,β)

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

$\mu = \alpha\beta, \ \sigma^2 = \alpha\beta^2$

$M(t) = (1-\beta t)^{-\alpha}, \ t < 1/\beta.$

or $\Gamma(v, \lambda)$

$f = \begin{cases} \frac{1}{\Gamma(v)} \lambda^v t^{v-1} e^{-\lambda t} & t>0 \\ 0 & o/w \end{cases}$

$F = \int f dx.$

$\mu = \frac{v}{\lambda}, \ \sigma^2 = \frac{v}{\lambda^2}, \ M(t) = \left(1 - \frac{t}{\lambda}\right)^{-v}$

⑤ Beta(α,β)

$f = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad x \in (0,1)$

$F = \int f$

$\mu = \frac{\alpha}{\alpha+\beta}, \ \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$

$M(t) = 1 + \sum_{i=1}^{\infty} \left(\prod_{j=0}^{j=k-1} \frac{\alpha+j}{\alpha+\beta+j}\right) \frac{t^i}{i!}, \ -\infty<t<\infty$

$\Rightarrow B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}$

⑥ Cauchy (a, λ) ⇒ Cauchy (0,1) = standard

$f = \frac{\lambda}{\pi(\lambda^2 + (t-a)^2)}$

$F = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{t-a}{\lambda}\right)$

✪ $\chi^2(r)$ $\Updownarrow$ $\Gamma(r/2, 2)$

$f(x) = \frac{1}{\Gamma(r/2) 2^{r/2}} x^{(r/2)-1} e^{-(x/2)}, \ x>0$

$\mu = r, \ \sigma^2 = 2r$

$M(t) = (1-2t)^{-r/2}, \ t < 1/2$

✪ T-dist: t(r): $f(x) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\,\Gamma(r/2)} \frac{1}{\left(1+\frac{x^2}{r}\right)^{(r+1)/2}}$, $-\infty < x < \infty$, $x \in (-\infty, \infty)$

If $r>1 \Rightarrow \mu = 0$, if $r>2$; $\sigma^2 = \frac{r}{r-2}$

✪ F, F(r₁,r₂): $r_1>0$, $r_2>0$

$f(x) = \frac{\Gamma\left(\frac{r_1+r_2}{2}\right)(r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{(x)^{r_1/2-1}}{\left(1+\frac{r_1 x}{r_2}\right)^{(r_1+r_2)/2}}, \ x>0$

If: $r_2>2$, $\mu = \frac{r_2}{r_2-2}$, if $r>4$, $\sigma^2 = 2\left(\frac{r_2}{r_2-2}\right)^2 \frac{r_1+r_2-2}{r_1(r_2-4)}$

✪ Logistic ~~Laplace~~: $-\infty<\theta<\infty$

$f = \frac{\exp\{-(x-\theta)\}}{(1+\exp\{-(x-\theta)\})^2} \quad -\infty<x<\infty$

$\mu = \theta, \ \sigma^2 = \pi^2/3$

$M(t) = e^{t\theta}\Gamma(1-t)\Gamma(1+t), \ -1<t<1$

✪ ~~Logistic~~, **Laplace** $-\infty<\theta<\infty$:

$f = \frac{1}{2} e^{-|x-\theta|} \quad -\infty<x<\infty$

$\mu = \theta, \ \sigma^2 = 2, \ M(t) = e^{t\theta}\frac{1}{1-t^2} \quad -1<t<1$

✪ Order Statistics pdf of k^{th} order statistic

$g_k(y_k) = \begin{cases} \frac{n!}{(k-1)!(n-k)!} [F(y_k)]^{k-1}[1-F(y_k)]^{n-k} f(y_k), & a<y_k<b \\ 0 & o/w \end{cases}$

Joint pdf of $(i,j)^{th}$ order statistic

$g_{ij}(y_i, y_j) =$