Wednesday 1/17

Grading

2 midterms	25% each
Final	35%
Quizzes	13% every 2 weeks on Fridays
HW	2% due every Friday on Canvas

1.1 Properties of Prob.

$(S, \mathcal{F}, P)$ is called a prob. space

Def → The set of all possible outcomes of an experiment is called a sample space (denoted by S).

$A \subseteq S$ is an event $\longrightarrow P(A)$ is called the prob. that A occurs

$\mathcal{F}$ = the collection of all events

A prob. measure $P: \mathcal{F} \longrightarrow \mathbb{R}$ satisfies

(i) $P(A) \geq 0$ for all $A \in \mathcal{F}$

(ii) $P(S) = 1$

(iii) If $A_1, A_2 \ldots$ are events and $A_i \cap A_j = \emptyset$ for all $i \neq j$ then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

In particular, if $A_1, A_2 \ldots A_K$ are events and $A_i \cap A_j = \emptyset$ for all $i \neq j$ then

$$P\left(\bigcup_{i=1}^{K} A_i\right) = \sum_{i=1}^{K} P(A_i)$$

e.g: Toss a coin $S = \{H, T\}$ $\quad P(\{H\}) = P(\{T\}) = \frac{1}{2} \quad P(\emptyset) = 0$

$\mathcal{F} = \{\{H\}, \{T\}, \{H, T\}, \emptyset\} \quad P(\{H, T\}) = 1$

Let A, B be events

$A \cup B$ = event that A or B occurs

$A \cap B$ = event that A and B occur

A^c = event that A does not occur

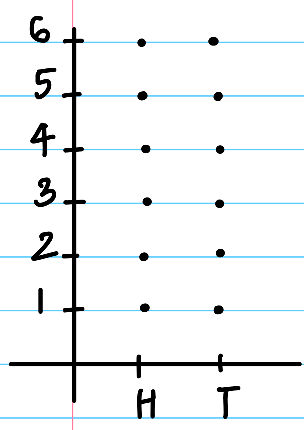
$A \setminus B = A \cap B^c$ = event that A occurs but not B

e.g: Toss a coin and then roll a die

$S = \{(H,1),(H,2)\ldots(H,6),(T,1),(T,2)\ldots(T,6)\}$

$= \{H,T\} \times \{1,2,3,4,5,6\}$ → Cartesian product

$A \times B = \{(i,j) \mid i \in A, j \in B\}$



→ total $2 \times 6 = 12$ pairs

Theorem →

1.1 – For any event A, $P(A^c) = 1 - P(A)$

$1 = P(S) = \underbrace{P(A \cup A^c)}_{\text{disjoint union}} = P(A) + P(A^c)$ by axiom (iii) $\therefore P(A^c) = 1 - P(A)$

1.2 – $P(\emptyset) = 0$

$P(\emptyset) = P(S^c) = 1 - P(S) = 1 - 1 = 0$

1.3 – If $A \subseteq B$ then $P(A) \leq P(B)$

$P(B) = P(A \cup (B \setminus A))$ disjoint union

$= P(A) + P(B \setminus A) \geq P(A) + 0$ by axiom (i) $\therefore P(B) \geq P(A)$

1.4 – For any event A, $P(A) \leq 1$

$A \subseteq S \therefore P(A) \leq P(S) = 1$

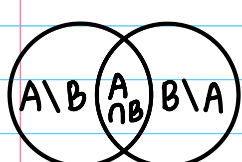
1.5 – Inclusion – Exclusion Principle $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \cup B) = P((A \setminus B) \cup (A \cap B) \cup (B \setminus A))$

$= P(A \setminus B) + P(A \cap B) + P(B \setminus A) = LHS$

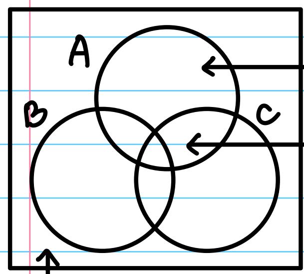
$P(A) + P(B) - P(A \cap B) = P(A \setminus B) + P(A \cap B) + P(B \setminus A) + \cancel{P(A \cap B)} - \cancel{P(A \cap B)}$

$= P(A \setminus B) + P(A \cap B) + P(B \setminus A) = RHS$



→ Total # sets to expand $2^n = 2^2 = 4$

1.6 – $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$



$A \cap B^c \cap C^c$

$A \cap B^c \cap C$

Could use either A or $A^c \rightarrow 2^n = 2^3 = 8$

$A^c \cap B^c \cap C^c$

Friday 1/19

$\boxed{1.1.9}$ $P(A_1) = \frac{2}{6} = \frac{1}{3}$ $P(A_2) = \frac{2}{6} = \frac{1}{3}$ $P(A_3) = \frac{2}{6} = \frac{1}{3}$

$i \neq j$ $P(A_i \cap A_j) = \frac{2.2}{6.6} = \frac{1}{9}$

$P(A_1 \cap A_2 \cap A_3) = \frac{2.2.2}{6.6.6} = \frac{1}{27}$

(b) Show that $P(A_1 \cup A_2 \cup A_3) = 1 - (1 - \frac{1}{3})^3$

$P(A_1 \cup A_2 \cup A_3) = 1 - P((A_1 \cup A_2 \cup A_3)^c)$

$= 1 - P(A_1^c \cap A_2^c \cap A_3^c)$ De Morgan's Law

$= 1 - \frac{4.4.4}{6.6.6} = 1 - \left(\frac{2}{3}\right)^3 = 1 - \left(1 - \frac{1}{3}\right)^3$

$\boxed{1.1.15}$ $S = A_1 \cup A_2 \ldots \cup A_m$

$A_1 \ldots A_m$ are mutually exclusive & exhaustive

mutually exclusive : $A_i \cap A_j = \emptyset$ for all $i \neq j$

exhaustive : $\bigcup_{i=1}^{m} A_i = S$

(a) Suppose $P(A_1) = \ldots = P(A_m)$ Show that $P(A_i) = \frac{1}{m}$ for all $i = 1, 2 \ldots m$

(b) If $A = A_1 \cup \ldots \cup A_n$ where $n < m$ then $P(A) = \frac{n}{m}$

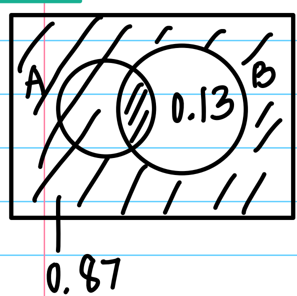
(a) $1 = P(S) = P(A_1 \cup \ldots \cup A_m) = \sum_{i=1}^{m} P(A_i) = \sum_{i=1}^{m} P(A_1) = m.P(A_1) \therefore P(A_1) = \frac{1}{m}$

$\therefore P(A_i) = P(A_1) = \frac{1}{m}$

(b) $A = A_1 \cup \ldots \cup A_h \quad (h < m)$

$$P(A) = \sum_{i=1}^{h} P(A_i) = \sum_{i=1}^{h} \frac{1}{m} = \frac{h}{m}$$

$\boxed{1.1.7}$ $P(A \cup B) = 0.76 \quad P(A \cup B') = 0.87$ Find $P(A)$



$P(A) = P(A \cup B) - P(B \setminus A)$
$= 0.76 - 0.13 = 0.63$

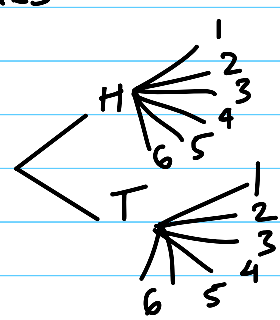
1.2 Methods of Enumeration

Multiplication Principle:

Suppose that an experiment E_1 has n_1 outcomes and for each of these possible outcomes, an experiment E_2 has n_2 possible outcomes. Then the composite experiment E_1E_2 that consists of performing first E_1 and then E_2 has n_1n_2 possible outcomes

e.g. Toss a coin and then roll a die

$\uparrow E_1$ (coin) $\quad \uparrow E_2$ (die)



$n_1 \; n_2$
$2.6 = 12$

Sampling theory →

Population size $= n$ Sample size $= r$

	Sampling w/o replacement	Sampling w/ replacement
Ordered sample = Permutation	$\underbrace{(n)(n-1)(n-2)\ \ldots\ (\ )}_{r}$, $n-(r-1) = n-r+1$ $nPr = n(n-1)\ldots(n-r+1)$ $= \frac{n!}{(n-r)!}$	$\underbrace{(n)(n)(n)\ \ldots\ }_{r} = n^r$

$nPn = \frac{n!}{0!} = n!$

Unordered sample = Combination

$nCr = \frac{n!}{(n-r)!\,r!} = \binom{n}{r}$

$\binom{n+r-1}{n-1} = \binom{n+r-1}{r}$

$n = 3$ Population $\{a,b,c\}$ $(a,b) \neq (b,a)$
$r = 2$ Choose 2

3 combinations nCr	6 permutations nPr
$\{a,b\}$	(a,b) (b,a)
$\{a,c\}$	(a,c) (c,a)
$\{b,c\}$	(b,c) (c,b)

$3C_2 \cdot 2 = 3P_2$

$\rightarrow 3C_2 = \frac{3P_2}{2}$

e.g: $n = 3$ $\{a,b,c\}$ = population set
Choose 5 w/ replacement $\rightarrow \{a,b,a,b,c\} = \{a,a,b,b,c\} \neq \{a,b,b,b,c\}$

$n_a + n_b + n_c = r$
$0 \leq n_a \leq r$
$0 \leq n_b \leq r$
$0 \leq n_c \leq r$

e.g: A deck of 52 cards. Choose 5 cards hand w/o replacement

(a) $P(\text{all 5 are spades}) = \frac{\binom{13}{5}}{\binom{52}{5}}$

(b) $P(\text{exactly 3 kings \& 2 queens}) = \frac{\binom{4}{3}\binom{4}{2}}{\binom{52}{5}}$

(c) $P(\text{2 kings, 2 queens, 1 jack}) = \frac{\binom{4}{2}\binom{4}{2}\binom{4}{1}}{\binom{52}{5}}$

Monday 1/22

Binomial Theorem $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

e.g: $(a+b)^3 = (a+b)(a+b)(a+b) = \binom{3}{2} a^2 b \ldots$

Proof → $(a+b)^n = (a+b)(a+b)\dots(a+b) = \sum x_1 x_2 \dots x_n$, $x_1 = a, b$ $x_2 = a, b \dots$

$$= \sum_{k=0}^{n} \sum a^k b^{n-k} \text{, exactly } k \text{ } x\text{'s are } a\text{'s} = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

To show the Pascal's triangle is true, show $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

Let $S = \{1, 2, 3 \dots n\}$ choose $k \rightarrow \binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

2 possibilities
- 1 already chosen → choose $(k-1)$ elements from remains
- 1 not chosen → choose k elements from remains

Multinomial Theorem

$$(a_1 + a_2 + a_3 + \dots + a_s)^n = \sum_{\substack{0 \le n_1, n_2 \dots n_s \le n \\ n_1 + n_2 + \dots + n_s = n}} \binom{n}{n_1}\binom{n-n_1}{n_2}\dots\binom{n-n_1-\dots n_{s-1}}{n_s} a_1^{n_1} a_2^{n_2} \dots a_s^{n_s}$$

Simplify coeffs:

$$= \frac{n!}{n_1!\,\cancel{(n-n_1)!}} \times \frac{\cancel{(n-n_1)!}}{n_2!\,\cancel{(n-n_1-n_2)!}} \times \dots \times \frac{\cancel{(n-n_1-n_2-\dots} - n_s)!}{n_s!\,\underbrace{(n-n_1-\dots-n_s)!}_{0! = 1}}$$

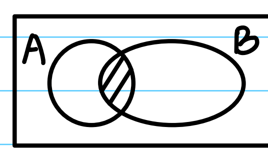
Notation

$$= \frac{n!}{n_1!\, n_2! \dots n_s!} = \binom{n}{n_1, n_2 \dots n_s}$$

1.3 Conditional Prob.

Def → The conditional prob. of A given that B has occurred is defined by

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



↳ Multiplication Rule: $\boxed{P(A \cap B) = P(B).\,P(A \mid B) = P(A).\,P(B \mid A)}$

e.g.: $\boxed{\begin{matrix} 7b \\ 3r \end{matrix}}$ Choose 2 w/o replacement

(a) $P(1^{st}\ red) = \frac{3}{10}$

(b) $P(1^{st}\ red, 2^{nd}\ blue) = P(1^{st}\ red) . P(2^{nd}\ blue \mid 1^{st}\ red) = \frac{3}{10} \times \frac{7}{9}$

(c) $P(1^{st}\ red, 2^{nd}\ red) = P(1^{st}\ red) . P(2^{nd}\ red \mid 1^{st}\ red) = \frac{3}{10} \times \frac{2}{9}$

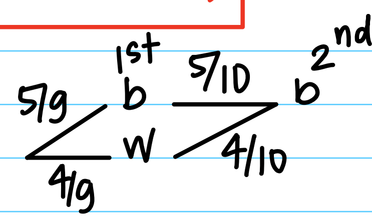
OR $\frac{\binom{3}{2}}{\binom{10}{2}}$

(d) $P(1\ red, 1\ blue) = \frac{\binom{3}{1}\binom{7}{1}}{\binom{10}{2}} = P(1^{st}\ red, 2^{nd}\ blue) + P(1^{st}\ blue, 2^{nd}\ red)$

3 events $\boxed{P(A \cap B \cap C) = P(A) . P(B \mid A) . P(C \mid A \cap B)}$

Wednesday 1/24

$\boxed{1.3.9}$ Box I 5b 4W Box II 4b 5W



Suppose we transfer 1 ball from Box I to Box II then choose 1 ball from Box II. What is the prob. that the sample ball is blue ?

$P(2^{nd}\ b) = P(1^{st}\ b \cap 2^{nd}\ b) + P(1^{st}\ w \cap 2^{nd}\ b)$

$= P(1^{st}\ b) P(2^{nd}\ b \mid 1^{st}\ b) + P(1^{st}\ w) P(2^{nd}\ b \mid 1^{st}\ w)$

$= \frac{5}{9} \times \frac{5}{10} + \frac{4}{9} \times \frac{4}{10} = \frac{41}{90}$

Theorem total prob → Let $\{B_1, B_2 \ldots B_n\}$ be a partition of the sample space S then

$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_n)$

$= P(B_1) . P(A \mid B_1) + P(B_2) . P(A \mid B_2) + \ldots + P(B_n) . P(A \mid B_n)$

We say that $\{B_1 \ldots B_n\}$ is a partition of S if $\bigcup_{i=1}^{n} B_i = S$ & $B_1 \ldots B_n$ are mutually exclusive

$\boxed{1.3.9}$ $\boxed{\begin{matrix} 1 & 2 \\ 3 & 4 \end{matrix}}$ Select balls one at a time w/o replacement

A_i = a match on the i^{th} draw, $i = 1, 2, 3, 4$

(a) Show $P(A_i) = \frac{3!}{4!}$ for each i

$\bigcirc\ \bigcirc\ \underset{\substack{\uparrow \\ \text{ith}}}{\textcircled{i}}\ \bigcirc$ $\quad P(A_i) = \frac{3!}{4!}$ ← permutations of other balls not i / ← permutations of 4 balls

(b) $P(A_i \cap A_j) = \frac{2!}{4!}$ for $i \neq j$ $\qquad P(A_i \cap A_j \cap A_k) = \frac{1!}{4!}$

2 fixed positions $\longrightarrow \frac{2!}{4!}$ $\qquad P(A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{1}{4!}$

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = P(A_1) + P(A_2) + P(A_3) + P(A_4) - [P(A_1 A_2) + \ldots] + [P(A_1 \cap A_2 \cap A_3) + \ldots] - P(A_1 \cap A_2 \cap A_3 \cap A_4)$$

$$= 4 \cdot \frac{3!}{4!} - \binom{4}{2}\frac{2!}{4!} + \binom{4}{3}\frac{1!}{4!} - \binom{4}{4}\frac{1}{4!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!}$$

Generally → n sets $A_1, A_2 \ldots A_n$

$$P(A_1 \cup A_2 \ldots \cup A_n) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots + (-1)^{n-1}\frac{1}{n!}$$

$$P(A_1^c \cap A_2^c \ldots \cap A_n^c) = 1 - P(A_1 \cup A_2 \ldots \cup A_n) = 1 - \frac{1}{1!} + \frac{1}{2!} - \ldots + (-1)^n \frac{1}{n!}$$

$$\xrightarrow{n \to \infty} e^{-1} = \frac{1}{e} = 0.37 \quad \left(e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!},\ x = -1 \right)$$

$\boxed{1.3.7}$ $\boxed{\begin{matrix} 2\ o \\ 2\ b \end{matrix}}$ Choose 2 balls w/o replacement

$$P(\text{both orange} \mid \text{at least 1 is orange}) = \frac{P(\text{both orange} \cap \text{at least 1 orange})}{P(\text{at least 1 orange})}$$

$$= \frac{P(\text{both orange})}{P(\text{at least 1 orange})} = \frac{\frac{\binom{2}{2}}{\binom{4}{2}}}{1 - P(2\text{ blues})} = \frac{\frac{1}{6}}{1 - \frac{1}{\binom{4}{2}}} = \frac{\frac{1}{6}}{1 - \frac{1}{6}} = \frac{1}{5}$$

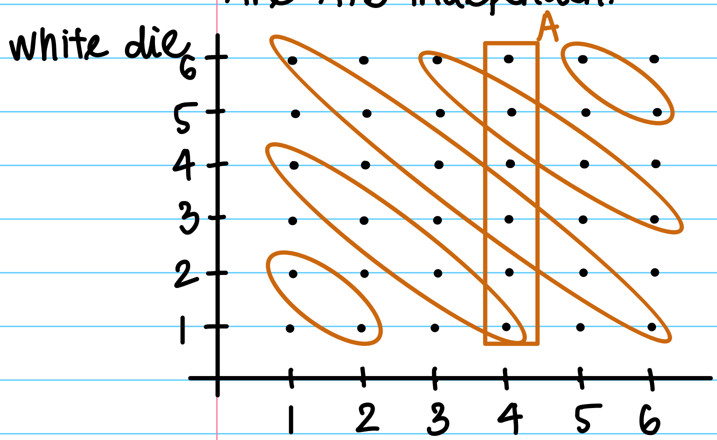
Monday
1/29

1.4 Independent Events

Def → Events A, B are independent if $\boxed{P(A \cap B) = P(A) . P(B)}$. Otherwise, A and B are called dependent events

Suppose $P(A) \neq 0$. Then $\boxed{A, B \text{ are independent} \iff P(B|A) = P(B)}$

e.g.: A red die and a white die are rolled.
Let $A = \{4 \text{ on the red die}\}$, $B = \{\text{sum of dice is odd}\}$
Are A, B independent?



red die

$P(A) = \frac{6}{36} = \frac{1}{6}$

$P(B) = \frac{18}{36} = \frac{1}{2}$

$P(A \cap B) = \frac{3}{36} = \frac{1}{12} = P(A) . P(B)$

OR $P(B|A) = \frac{3}{6} = \frac{1}{2} = P(B)$

$\therefore A, B$ independent

Theorem → $\boxed{A, B \text{ independent} \iff \begin{cases} A^c, B^c \\ A, B^c \text{ independent} \\ A^c, B \end{cases}}$

Proof → $P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A) . P(B) = P(A) . (1 - P(B)) = P(A) . P(B^c)$
$\therefore A, B^c$ independent

Def → A, B, C are mutually independent if
(i) $P(A \cap B) = P(A) . P(B)$
(ii) $P(A \cap C) = P(A) . P(C)$
(iii) $P(B \cap C) = P(B) . P(C)$
} pairwise independent
and $P(A \cap B \cap C) = P(A) . P(B) . P(C)$

n sets → $A_1, A_2 \ldots A_n$ are mutually independent if
$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k})$ for all $k = 1, 2 \ldots n$
and all $1 \leq i_1 < i_2 \ldots < i_k \leq n$

General $\rightarrow$ $\boxed{\begin{array}{l} A_1, A_2 \ldots A_n \text{ are mutually independent if} \\ P(A_{i_1}^* \cap A_{i_2}^* \cap \ldots \cap A_{i_k}^*) = P(A_{i_1}^*) \ldots P(A_{i_k}^*) \text{ where } A^* = A \text{ or } A^c \end{array}}$

e.g: A fair 6-sided is rolled 6 times. Let A_i be the event that side i is observed on the i^{th} roll, called a match on the i^{th} trial, $i = 1, 2 \ldots 6$

Find $P(\text{no matches occur}) = P(A_1^c \cap A_2^c \ldots \cap A_6^c)$

$$= P(A_1^c) . P(A_2^c) \ldots P(A_6)^c = \left(\frac{5}{6}\right)^6$$

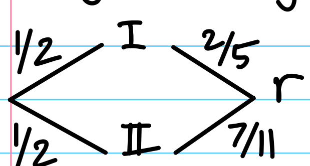
1.5 Bayes' Theorem

Theorem $\rightarrow$ Let $\{B_1, B_2 \ldots B_n\}$ be a partition of the sample space. Then:

$$\boxed{P(B_i \mid A) = \frac{P(B_i) . P(A \mid B_i)}{P(A)} = \frac{P(B_i) . P(A \mid B_i)}{\sum_{j=1}^{n} P(B_j) . P(A \mid B_j)}}$$

$P(B_i)$ $i = 1, 2 \ldots n$ are called prior probs and $P(B_i \mid A)$ $i = 1, 2 \ldots n$ are called the posterior probs.

e.g: Box I — 2r, 3g; Box II — 7r, 4g. Choose a box at random and draw a ball



$$P(I \mid r) = \frac{P(I \cap r)}{P(r)} = \frac{P(I) . P(r \mid I)}{P(I) . P(r \mid I) + P(II) . P(II \mid r)}$$

$$= \frac{1/2 \times 2/5}{1/2 \times 2/5 + 1/2 \times 7/11} = \frac{22}{57}$$

Wednesday 1/31

2.1 Random Variables of the Discrete Type

Def $\rightarrow$ Let $(S, \mathcal{F}, P)$ be a prob. space. A real-valued fn defined on S is called a random variable $X : S \rightarrow \mathbb{R}$

S_X = range of $X = \{X(w) \mid w \in S\}$ is called the space of X

μ_X is a prob. measure on S_X, a prob. distribution of X

e.g: Toss a coin. Let $X = \#$ heads. $S = \{H, T\} \longrightarrow X(H) = 1,\ X(H) = 0$

$S_X = \{0, 1\} \subseteq \mathbb{R}$

Fair coin: $P(\{H\}) = P(\{T\}) = \frac{1}{2}$ $\to$ $\mu_X(1) = P(X=1) = P(\{H\}) = 1/2$, $\mu_X(0) = P(X=0) = P(\{T\}) = 1/2$

$(S, \mathcal{F}, P) \xrightarrow{X} (S_X, \mathcal{G}, \mu_X)$

$\mathcal{G}$: the class of all subsets of S_X

$\mathcal{F}$: the class of all subsets of S

e.g: $f_X(0) = \frac{1}{2},\ f_X(1) = \frac{1}{2} \longrightarrow$ $\boxed{f_X \text{ is the pmf of } X:\ f_X(i) = P(X=i) = \mu_X(\{i\})}$

Def $\rightarrow$ A random variable X is called discrete if S_X is finite or countably infinite

$S_X = \{x_1, x_2 \dots\}$

In this case, $f_X(x_i)$, $i = 1, 2 \dots$ are sufficient for the distribution of X

$$\boxed{f_X(x_i) = P(X = x_i) \text{ is the pmf of } X}$$

e.g: Roll a die 2 times. Let $X =$ maximum of the 2 outcomes

$S_X = \{1, 2, 3, 4, 5, 6\}$. Find pmf of X

$f_X(1) = \frac{1}{36},\ f_X(2) = \frac{3}{36},\ f_X(3) = \frac{5}{36},\ f_X(4) = \frac{7}{36},\ f_X(5) = \frac{9}{36},\ f_X(6) = \frac{11}{36}$

$\to$ definitely adds up to 1

2.2 Mathematical Expectation

Let X be a discrete random variable w/ pmf $f_X(x)$, $x \in S_X$ then $E[X]$ is defined by

$$\boxed{E[X] = \sum_{x \in S_X} x f_X(x)} \quad \text{weighted sum}$$

Def $\rightarrow$

$$\boxed{E[g(X)] = \sum_{x \in S_X} g(x) f_X(x)}$$

e.g: Let X be a random variable w/ pmf

x	0	1	2	3
f_X	.2	.3	.1	.4

Let $Y = X^2$. Find $E[Y]$

① $E[Y] = E[X^2] = \sum_{i=0}^{3} i^2 f_X(i) = 0^2 \times 0.2 + 1^2 \times 0.3 + 2^2 \times 0.1 + 3^2 \times 0.4 = 4.3$

②

y	0	1	4	9
f_Y	.2	.3	.1	.4

$\rightarrow E[Y] = 0 \times 0.2 + 1 \times 0.3 + 4 \times 0.1 + 9 \times 0.4 = 4.3$

(1) $E[f(X) + g(X)] = E[f(X)] + E[g(X)]$
(2) $E[af(X)] = aE[f(X)]$
(3) $E[c] = c$

Let X be a discrete random variable, $\mu = E[X]$ then μ minimizes $E[(X-b)^2]$ over $b \in \mathbb{R} \longrightarrow \mu = \text{argmin}\, E[(X-b)^2],\ b \in \mathbb{R}$

Proof →

Let $h(b) = E[(X-b)^2]$

$h'(b) = E[2(X-b)] = 0 \rightarrow E[(X-b)] = 0 = E[X] - \underbrace{E[b]}_{b} \rightarrow E[X] = b$

$\therefore h(b)$ has a min at $b = \mu$ $\therefore \mu = \text{argmin}\, E[(X-b)^2], b \in \mathbb{R}$ Q.E.D!

Friday 2/2

2.3 Special Mathematical Expectations

Let X be a discrete random variable w/ pmf $f(x)$, $x \in S_X$

Def →

$$\boxed{E[g(X)] = \sum_{x \in S_X} g(x) f_X(x)}$$

Mean of $X = \mu = E[X] = \sum_x x f_X(x)$

Variance of $X = Var(X) = \sigma^2 = E[(X-\mu)^2]$

e.g: $X = \begin{cases} 1, & P = 1/2 \\ -1, & P = 1/2 \end{cases}$ $\mu_X = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0$

$\sigma_X^2 = E[(X - \underbrace{\mu}_{0})^2] = E[X^2] = 1^2 \times \frac{1}{2} + (-1)^2 \times \frac{1}{2} = 1$

e.g: $Y = \begin{cases} 10, & P = 1/2 \\ -10, & P = 1/2 \end{cases}$ $\mu_Y = 0$ $\sigma_Y^2 = 10^2 \times \frac{1}{2} + (-10)^2 \times \frac{1}{2} = 100$

Formula $\rightarrow$ $\boxed{\sigma^2 = E[(X-\mu)^2] = E[X^2] - \mu^2}$

Proof $\rightarrow$

$$E[(X-\mu)^2] = E[X^2 - 2X\mu + \mu^2] = E[X^2] - 2\underbrace{E[X]}_{\mu}\mu + E[\mu^2]$$
$$= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2 \quad \text{Q.E.D!}$$

e.g: Let $X \sim Geom(p)$, $0 < p \leq 1$ $\qquad f(x) = q^{k-1}p$, $k = 1, 2, 3 \ldots$
$q = 1-p$

(a) Find $E[X] = \mu_X$

$$\mu_X = \sum_{k=1}^{\infty} k f(k) = \sum_{k=1}^{\infty} k q^{k-1} p = p\sum_{k=1}^{\infty} k q^{k-1} = p\sum_{k=1}^{\infty} (q^k)'$$
$$= p\left(\sum_{k=1}^{\infty} q^k\right)' = p\left(\sum_{k=0}^{\infty} q^k\right)' = p\left(\frac{1}{1-q}\right)' = p((1-q)^{-1})'$$
$$= p(-1)(-1)(1-q)^{-2} = \frac{p}{p^2} = \frac{1}{p}$$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

(b) Find σ_X^2 $\longrightarrow$ consider 2nd factorial moment

$$E[X(X-1)] = \sum_{k=1}^{\infty} k(k-1) f(k) = \sum_{k=1}^{\infty} k(k-1) q^{k-1} p = pq\sum_{k=1}^{\infty} k(k-1) q^{k-2}$$
$$= pq\sum_{k=1}^{\infty} (q^k)'' = pq\left(\sum_{k=1}^{\infty} q^k\right)'' = pq\left(\sum_{k=0}^{\infty} q^k\right)'' = pq\left(\frac{1}{1-q}\right)''$$
$$= pq((1-q)^{-2})' = pq(-2)(-1)(1-q)^{-3} = \frac{2pq}{p^3} = \frac{2q}{p^2}$$

$$E[X^2] = E[X(X-1)] + E[X] = \frac{2q}{p^2} + \frac{1}{p}$$

$$\sigma_X^2 = E[X^2] - \mu_X^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q+p-1}{p^2} = \frac{2q-q}{p^2} = \frac{q}{p^2}$$

Def $\rightarrow$ $\boxed{M_X(t) = E[e^{tX}]}$ is called the moment generating function of X, $t \in \mathbb{R}$

$\boxed{E[X^n] = M_X^{(n)}(0)}$

Proof $\rightarrow$ $M_X'(t) = \frac{d}{dt}E[e^{tX}] = E[Xe^{tX}] \longrightarrow M_X''(t) = E[X^2e^{tX}]$

$\therefore M_X^{(n)}(t) = E[X^n e^{tX}] \qquad \therefore M_X^{(n)}(0) = E[X^n]$ Q.E.D!

(*) Geometric distribution X:

$$M_X(t) = E[e^{tX}] = \sum_{k=1}^{\infty} e^{tk} f(k) = \sum_{k=1}^{\infty} e^{tk} q^{k-1} p = pe^t \sum_{k=1}^{\infty} e^{t(k-1)} q^{k-1}$$

$$= pe^t \sum_{k=1}^{\infty} (e^t q)^{k-1} \overset{s=k-1}{=\!=} pe^t \sum_{s=0}^{\infty} (e^t q)^s = \frac{pe^t}{1-qe^t}$$

$$qe^t < 1 \rightarrow e^t < \frac{1}{q} \rightarrow t < \ln\left(\frac{1}{q}\right) = -\ln q = -\ln(1-p)$$

Monday 2/5

$$M(t) = pe^t(1-qe^t)^{-1}$$

$$M'(t) = pe^t(1-qe^t)^{-1} + pe^t(-qe^t)(-1)(1-qe^t)^{-2}$$

$$= pe^t(1-qe^t)^{-1} + pqe^{2t}(1-qe^t)^{-2}$$

$$M''(t) = pe^t(1-qe^t)^{-1} + pqe^{2t}(1-qe^t)^{-2} + 2pqe^t(1-qe^t)^{-2} + 2pq^2e^{3t}(1-qe^t)^{-3}$$

$$= pe^t(1-qe^t)^{-1} + 3pqe^{2t}(1-qe^t)^{-2} + 2pq^2e^{3t}(1-qe^t)^{-3}$$

$$M'''(t) = pe^t(1-qe^t)^{-1} + pqe^{2t}(1-qe^t)^{-2} + 6pqe^t(1-qe^t)^{-2} + 3pq^2e^{3t}\cdot 2(1-qe^t)^{-3}(qe^t) + 2pq^2e^{3t}\cdot 3(1-qe^t)^{-3} + 2pq^2e^{3t}\cdot 3(1-qe^t)^{-4}(qe^t)$$

$$\therefore E[X^3] = M'''(0) = p\cdot\frac{1}{p} + \frac{pq}{p^2} + 6pq\cdot\frac{1}{p^2} + \frac{6pq^2}{p^3} + \frac{6pq^2}{p^3} + \frac{6pq^3}{p^4}$$

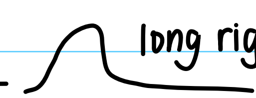
$$= 1 + \frac{q}{p} + \frac{6q}{p} + \frac{6q^2}{p^2} + \frac{6q^2}{p^2} + \frac{6q^3}{p^3}$$

$$= \frac{1}{p} + \frac{6q}{p}\left(1 + \frac{2q}{p}\right) + \frac{6q^3}{p^3}$$

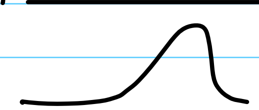
$$= \frac{1}{p} + \frac{6q}{p} \times \frac{q+1}{p} + \frac{6q^3}{p^3} = \frac{1}{p} + \frac{6q}{p^2}\left(q + 1 + \frac{q^2}{p}\right)$$

$$\boxed{\gamma = \text{skewness} = \frac{E[(X-\mu)^3]}{\sigma^3}} \qquad = E[X^3] - 3\mu\sigma^2 - \mu^3$$

$\gamma > 0$ skewed to right



$\gamma < 0$ skewed to left



$$\boxed{\text{Geom}(p) \longrightarrow \gamma = \frac{2-p}{\sqrt{1-p}}}$$

Theorem Let $g(t) = \ln M(t)$ then (1) $g'(0) = \mu$

(2) $g''(0) = E[(X-\mu)^2] = \sigma^2$

(3) $g'''(0) = E[(X-\mu)^3]$

$$\boxed{\gamma = \frac{g'''(0)}{\sigma^3}}$$

Proof

$$g'(t) = \frac{M'(t)}{M(t)} \longrightarrow g'(0) = \frac{M'(0)}{M(0)} = \frac{E[X]}{1} = \mu$$

e.g.: $X \sim$ Poisson$(\lambda) \longrightarrow f(x) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, 2 \ldots$

$$M_X(t) = E[e^{tX}] = \sum_{k=1}^{\infty} e^{tk}\frac{\lambda^k}{k!}e^{-\lambda} = \left(\sum_{k=1}^{\infty}\frac{(e^t\lambda)^k}{k!}\right)e^{-\lambda}$$

$$\sum_{k=0}^{\infty}\frac{x^k}{k!} = e^x$$

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

$g(t) = \ln M_X(t) = \lambda(e^t - 1)$

$\hookrightarrow g'(t) = \lambda e^t \longrightarrow \mu = g'(0) = \lambda$

$g''(t) = \lambda e^t \longrightarrow \sigma^2 = g''(0) = \lambda \longrightarrow \sigma = \sqrt{\lambda}$

$g'''(t) = \lambda e^t \longrightarrow E[(X-\mu)^3] = g'''(0) = \lambda \longrightarrow \gamma = \frac{g'''(0)}{\sigma^3} = \frac{\lambda}{(\sqrt{\lambda})^3} = \frac{1}{\sqrt{\lambda}}$

Wednesday 2/7

2.4 The Binomial Distribution

Def An experiment is called a Bernoulli trial if it has exactly 2 possible outcomes (success/failure) — $P(\text{success}) = p$, $P(\text{failure}) = q = 1-p$

Let X = #successes in 1 trial $\longrightarrow X = 0$ or 1

pmf:

X	0	1
f_X	q	p

$$\boxed{X \sim Ber(p)}$$

Consider a sequence of independent Bernoulli trials w/ success prob. p for each trial.

Let X = #successes in n trials $\longrightarrow$ pmf: $\boxed{f_X(k) = \binom{n}{k}p^k q^{n-k}}$ $k = 0, 1, 2 \ldots$

$X \sim Bin(n, p)$

$K = 1, 2 \ldots$

$N \sim Geom(p)$

Let N = # trials needed to get 1st success $\longrightarrow$ pmf: $\boxed{f_X(K) = q^{K-1}p}$

e.g.: Let $X \sim Bin(n, p)$

(a) Find μ_X (b) σ_X^2 (c) $M_X(t)$ (d) γ_X

(c) $M_X(t) = E[e^{tX}] = \sum_{K=0}^{n} e^{tK} \binom{n}{K} p^K q^{n-K} = \sum_{K=0}^{n} \binom{n}{K} (pe^t)^K q^{n-K}$

$= (pe^t + q)^n = (q + pe^t)^n$

↳ binomial theorem $(a+b)^n$

Use Log theorem $\longrightarrow$ Let $g(t) = \ln M(t) = \ln(q + e^t)^n = n \ln(q + pe^t)$

(a) $\mu_X = g'(0) = n . \frac{pe^t}{q + pe^t}\Big|_{t=0} = \frac{np}{q+p} = np$

(b) $\sigma_X^2 = g''(0) = n(pe^t(q+pe^t)^{-1})'|_{t=0}$

$= n(pe^t(q+pe^t)^{-1} + (pe^t)^2(-1)(q+pe^t)^{-2})|_{t=0}$

$= n(p\underbrace{(q+p)}_{1}{}^{-1} - p^2\underbrace{(q+p)}_{1}{}^{-2}) = npq \longrightarrow \sigma_X = \sqrt{npq}$

(d) $E[(X-\mu)^3] = g'''(0)$

$= n(pe^t(q+pe^t)^{-1} + (pe^t)(-1)(q+pe^t)^{-2}(pe^t) - 2p^2e^{2t}(q+pe^t)^{-2} - p^2e^{2t}(-2)(q+pe^t)^{-3}pe^t)|_{t=0}$

$= n(p - p^2 - 2p^2 + 2p^3) = np(1 - 3p + 2p^2) = np(\underbrace{1-p}_{q})(1-2p) = npq(1-2p)$

$\therefore \gamma_X = \frac{E[(X-\mu)^3]}{\sigma_X^3} = \frac{npq(1-2p)}{(npq)^{3/2}} = \frac{1-2p}{\sqrt{npq}}$,
$\gamma_X > 0 \Leftrightarrow p < 1/2$
$\gamma_X < 0 \Leftrightarrow p > 1/2$
$\gamma_X = 0 \Leftrightarrow p = 1/2$

W/o log $\longrightarrow$ $\mu_X = \sum_{K=0}^{n} \binom{n}{K} p^K q^{n-K} = \sum_{K=0}^{n} \frac{n!}{K!(n-K)!} p^K q^{n-K}$

$= \sum_{K=1}^{n} \frac{n!}{(n-K)!(K-1)!} p^K q^{n-K} = \sum_{s=0}^{n} np. \underbrace{\frac{(n-1)!}{s!(n-1-s)!} p^s q^{n-1-s}}_{Bin(n-1,\, s)} = np$

$s = K - 1 \rightarrow s + 1 = K$

Monday 2/12

2.5 The Hypergeometric Distribution

N_1 (A)
N_2 (B)

In a box, choose n w/o replacement

Let X = # of type (A) balls in the sample

then $X \sim HG(N_1, N_2, n)$

pmf of X: $X = 0, 1, 2 \dots \min\{N_1, n\}$

$$\boxed{f_X(k) = \frac{\binom{N_1}{k}\binom{N_2}{n-k}}{\binom{N}{n}}} \qquad N = N_1 + N_2$$

$$M_X(t) = E[e^{tX}] = \sum_{k=0}^{\min\{N_1, n\}} e^{tk} \frac{\binom{N_1}{k}\binom{N_2}{n-k}}{\binom{N_1+N_2}{n}}$$

$$\mu_X = E[X] = \sum_{k=0}^{N_1} k \frac{\binom{N_1}{k}\binom{N_2}{n-k}}{\binom{N}{n}} = \sum_{k=1}^{N_1} \frac{\frac{N_1!}{(k-1)!(N_1-k)!} \times \frac{N_2!}{(n-k)!(N_2-n+k)!}}{\frac{N!}{n!(N-n)!}}$$

$k=0 \rightarrow \mu_X + = 0 \rightarrow$ cancel k & change bounds

$s = k-1$

$k = s+1$

$$= N_1 \sum_{s=0}^{N_1-1} \frac{\frac{(N_1-1)!}{s!(N_1-1-s)!} \times \frac{N_2!}{(n-1-(k-1))!(N_2-(n-1)+(k-1))!}}{\frac{N!}{((N-1)-(n-1))!(n-1)!\,n}}$$

$$= \frac{nN_1}{N} \underbrace{\sum_{s=0}^{N_1-1} \frac{\frac{(N_1-1)!}{s!(N_1-1-s)!} \times \frac{N_2!}{(n-1-s)!(N_2-(n-1)+s)!}}{\frac{(N-1)!}{((N-1)-(n-1))!(n-1)!}}}_{HG(N_1-1,\, N_2,\, n-1)} \qquad \therefore \boxed{\mu_X = \frac{nN_1}{N}}$$

similar

e.g.: $X \sim Bin(n, \frac{N_1}{N})$ or $Bin(n, p) \rightarrow \mu_X = np$

$\sigma^2 = npq = n.\frac{N_1}{N}.\frac{N_2}{N}$ (w/ replacement)

Find σ_X^2, for $X \sim HG(N_1, N_2, n)$

$$E[X^2] = E[X(X-1)] + E[X]$$

$$E[X(X-1)] = \sum_{k=0}^{N_1} \cancel{k(k-1)} \frac{\frac{N_1!}{(N_1-k)!\cancel{k!}} \times \frac{N_2!}{(N_2-n+k)!(n-k)!}}{\frac{N!}{(N-n)!\,n!}}$$

$$= \sum_{k=2}^{N_1} \frac{\frac{N_1!}{(N_1-k)!(k-2)!} \times \frac{N_2!}{(N_2-n+k)!(n-k)!}}{\frac{N!}{(N-n)!\,n!}}$$

$s = k-2$
$k = s+2$

$$= \sum_{s=0}^{N_1-2} \frac{\frac{N_1(N_1-1)(N_1-2)!}{(N_1-s-2)!\,s!} \times \frac{N_2!}{(N_2-(n-2)+s)!(n-2-s)!}}{\frac{N(N-1)(N-2)!}{(N-2-(n-2))!(n-2)!\,n(n-1)}}$$

$$= n(n-1)\frac{N_1(N_1-1)}{N(N-1)} \underbrace{\sum_{s=0}^{N_1-2} \frac{\binom{N_1-2}{s}\binom{N_2}{n-2-s}}{\binom{N-2}{n-2}}}_{HG(N_1-2,\ N_2,\ n-2)} = n(n-1)\frac{N_1(N_1-1)}{N(N-1)}$$

$$\therefore \sigma^2 = E[X^2] - \mu_X^2 = E[X(X-1)] + \mu_X - \mu_X^2$$

$$= n(n-1)\frac{N_1(N_1-1)}{N(N-1)} + n\frac{N_1}{N} - n^2\frac{N_1^2}{N^2}$$

$$= n\frac{N_1}{N}\left(\frac{(n-1)(N_1-1)}{N-1} + 1 - n\frac{N_1}{N}\right)$$

$$= n\frac{N_1}{N} \times \frac{N(n-1)(N_1-1) + N(N-1) - nN_1(N-1)}{N(N-1)}$$

$$= n\frac{N_1}{N} \times \frac{\cancel{nN_1N} - nN - N_1N + \cancel{N} + N^2 - \cancel{N} - \cancel{nN_1N} + nN_1}{N(N-1)}$$

$n=1: \sigma^2 = npq$
$n=n: \sigma^2 = 0$

$$= n\frac{N_1}{N} \times \frac{1}{N(N-1)}\left(\underbrace{n(N_1-N)}_{-nN_2} + \underbrace{N(N-N_1)}_{NN_2}\right) = \boxed{n\frac{N_1}{N}\frac{N_2}{N}\cdot\frac{N-n}{N-1}}$$

$$-nN_2 + NN_2 = N_2(N-n)$$

Newton's Binomial Theorem:

$$\boxed{\frac{1}{(1-w)^r} = \sum_{k=0}^{\infty} \binom{r+k-1}{r-1} w^k} \qquad -1 < w < 1$$

Proof $\longrightarrow$

$$(b+1)^n = \sum_{k=0}^{n} \binom{n}{k} b^k = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} b^k = \sum_{k=0}^{n} \frac{n(n-1)\dots(n-k+1)}{k!} b^k$$

$$\therefore (1-w)^{-r} = \sum_{k=0}^{n} \frac{(-r)(-r-1)\dots(-r-k+1)}{k!} (-w)^k$$

$$= \underbrace{(-1)^k (-1)^k}_{1^k = 1} \sum_{k=0}^{n} \frac{r(r+1)\dots(r+k-1)}{k!} w^k$$

$$= \sum_{k=0}^{n} \frac{(r+k-1)!}{(r-1)!\,k!} w^k = \sum_{k=0}^{n} \binom{r+k-1}{r-1} w^k \quad \text{Q.E.D!}$$

Wednesday 2/14 Let $s = k + r$ then $\boxed{\frac{1}{(1-w)^r} = \sum_{s=r}^{\infty} \binom{s-1}{r-1} w^{s-r}} \quad |w| < 1$

2.6 Negative Binomial Distribution

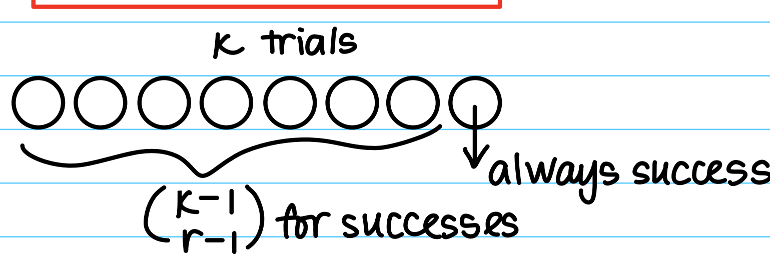
Consider a sequence of independent Bernoulli trials Ber(p)

Let X = # trials needed to obtain r successes

$r = 1 \longrightarrow X \sim \text{Geom}(p)$

For general, $r = 1, 2 \dots \longrightarrow X \sim \text{Negbin}(r, p)$

$X = r, r+1, r+2 \dots$ pmf: $\boxed{f_X(k) = \binom{k-1}{r-1} p^r q^{k-r}}$ $k = r, r+1, r+2 \dots$



(a) Verify that $f_X(k)$ is a pmf

$$\sum_{k=r}^{\infty} f_X(k) = \sum_{k=r}^{\infty} \binom{k-1}{r-1} p^r q^{k-r} = p^r \sum_{k=r}^{\infty} \binom{k-1}{r-1} q^{k-r}$$

↳ use 2nd binomial formula

$$= p^r \frac{1}{(1-q)^r} = \frac{p^r}{p^r} = 1$$

(b) Find $M_X(t) = \sum_{k=r}^{\infty} e^{tk} \binom{k-1}{r-1} p^r q^{k-r}$

$$= p^r \sum_{k=r}^{\infty} e^{(k-r)t} e^{rt} \binom{k-1}{r-1} q^{k-r}$$

$$= (pe^t)^r \sum_{k=r}^{\infty} \binom{k-1}{r-1} (qe^t)^{k-r} = (pe^t)^r \frac{1}{(1-qe^t)^r} = \boxed{\left(\frac{pe^t}{1-qe^t}\right)^r}$$

Summary → $M_X(t)$

$Ber(p)$	$q+pe^t$	$Y \sim Geom(p)$	$\dfrac{pe^t}{1-qe^t}$
$Bin(n,p)$	$(q+pe^t)^n$	$X \sim Negbin(r,p)$	$\left(\dfrac{pe^t}{1-qe^t}\right)^r$

(c) Find $\mu_X \longrightarrow g_X(t) = \ln M_X(t) = r\ln\left(\frac{qe^t}{1-pe^t}\right) = r\ln M_Y(t) = rg_Y(t)$

$$g_X'(t) = rg_Y'(t) \xrightarrow{t=0} \mu_X = r\mu_Y = r \times \frac{1}{p} = \frac{r}{p}$$

$$g_X''(t) = rg_Y''(t) \xrightarrow{t=0} \sigma_X^2 = r\sigma_Y^2 = r \times \frac{q}{p^2} = \frac{rq}{p^2}$$

$$g_X'''(t) = rg_Y'''(t) \xrightarrow{t=0} \gamma_X = \frac{rg_Y'''(0)}{\sigma_X^3} = \frac{rg_Y'''(0)}{\left(\frac{rq}{p^2}\right)^{3/2}} = \frac{1}{\sqrt{r}}\gamma_Y$$

Friday 2/16

Taylor Expansion: $M_X(t) = f(t) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} t^k \longrightarrow E[X^k] = M_X^{(k)}(0)$

2.7 The Poisson Distribution

$X \sim Poisson(\lambda)$ if pmf: $\boxed{f_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}}$, $k = 0, 1, 2 \ldots$ $\lambda > 0$

$\mu_X = \lambda \quad \sigma_X^2 = \lambda \quad \gamma_X = \frac{1}{\sqrt{\lambda}} \quad M_X(t) = e^{\lambda(e^t - 1)}, \ t \in \mathbb{R}$

Theorem → $\boxed{\text{Let } S_n \sim Bin(n,p) \ \& \ p = \frac{\lambda}{n} \text{ then } S_n \longrightarrow X \sim Poisson(\lambda), \ n \to \infty}$

e.g.: $S_n \sim Bin(40, 0.05)$. Find $P(S_n = 3) = \binom{40}{3} 0.05^3 \, 0.95^{37} = 0.1851$

OR $E[S_n] = np = 40(0.05) = 2 \longrightarrow S_n \sim Bin(40, 0.05) \approx Poisson(np) = Poisson(2)$

$P(S_n = 3) \sim P(X = 3) = \frac{2^3}{3!} e^{-2} = 0.1804$ ← computing

table → $= P(X \leq 3) - P(X \leq 2) = 0.857 - 0.677 = 0.180$

Proof →

$$P(S_n = k) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n!}{k!(n-k)!} \cdot \lambda^k \cdot \frac{1}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \frac{1}{\left(1 - \frac{\lambda}{n}\right)^k}$$

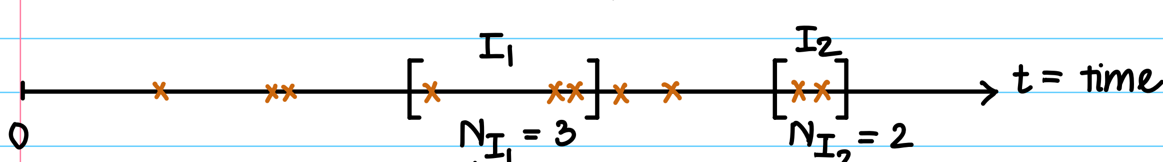
$$= \frac{\lambda^k}{k!}\left(1 - \frac{\lambda}{n}\right)^n \cdot \cancelto{1}{\frac{n}{n}} \times \cancelto{1}{\frac{n-1}{n}} \times \ldots \times \cancelto{1}{\frac{n-k+1}{n}} \cdot \cancelto{1}{\frac{1}{\left(1 - \frac{\lambda}{n}\right)^k}}$$

$$\overset{n \to \infty}{=} \frac{\lambda^k}{k!} e^{-\lambda}, \ k = 0, 1, 2 \ldots \quad \text{Q.E.D!}$$

Summary → $Bin(n,p) \approx \begin{cases} Poisson(np) & \text{if } np^2 < 0.05, \ n > 25 \\ N(np, npq) & \text{if } npq > 10, \ n > 25 \end{cases}$

Monday 2/19

* Poisson process → events that occur along an axis of time (phone calls)



↓

phone calls within interval I

Assume N_I is Poisson distribution

Def →

A family of random variables $\{N_I, I \subseteq [0,\infty)\}$ is called a Poisson process w/ intensity λ if

(1) $N_I \sim \text{Poisson}(\lambda . |I|)$ ↓ length I

(2) $N_{I_1}, N_{I_2} \dots N_{I_n}$ are independent if I_i and I_j are non-overlapping for all $i \neq j$

e.g: In a large city: avg all calls = 2/3 mins. If assume Poisson process, prob. of 5 or more calls in 9 mins period?

$$\lambda = \frac{2}{3} \longrightarrow N_{[0,9]} \sim \text{Poisson}(\lambda t) = \text{Poisson}\left(\frac{2}{3} \times 9 \text{ mins}\right) = \underset{\text{new } \lambda}{\text{Poisson}(6)}$$

$$P(N_{[0,9]} \geq 5) = 1 - P(N_{[0,9]} \leq 4) = 1 - \sum_{k=0}^{4} \frac{6^k}{k!} e^{-6} = 0.715$$

3.1 Random Continuous Variables

Def →

A random variable X is called continuous if there is a fn $f(x)$ that:

$$\boxed{P(X \in B) = \int_B f(x)\,dx, \text{ for all } B \subseteq \mathbb{R}} \quad \boxed{\text{cdf: } F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx}$$

↓

prob. density fn at X: pdf $f(x)$, $f_X(x)$

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx \qquad \mu_X = \int_{-\infty}^{\infty} x f_X(x)\,dx$$

$$\sigma_X^2 = E[(X-\mu_X)^2] = \int_{-\infty}^{\infty} (X-\mu_X)^2 f_X(x)\,dx = E[X^2] - \mu_X^2$$

e.g: pdf: $f(x) = xe^{-x}$, $0 \leq x \leq \infty$

(a) Find $M_X(t)$

$$M_X(t) = E[e^{tX}] = \int_0^{\infty} e^{tx} f_X(x)\,dx = \int_0^{\infty} e^{tx} x e^{-x}\,dx = \int_0^{\infty} x e^{(t-1)x}\,dx$$

$t<1$ $$= \int_0^\infty x e^{-(t-1)x}\,dx = x \cdot \frac{1}{-(1-t)} e^{-(1-t)x}\Big|_0^\infty - \int_0^\infty \frac{1}{-(1-t)} e^{-(1-t)x}\,dx$$

$$= \frac{1}{1-t}\int_0^\infty e^{-(1-t)x}\,dx = \frac{-1}{(1-t)^2} e^{-(1-t)x}\Big|_0^\infty = \frac{1}{(1-t)^2}$$

(b) $\mu_X = ?,\ \sigma_X^2 = ?$

$$g(t) = \ln M_X(t) = \ln \frac{1}{(1-t)^2} = -2\ln(1-t)$$

$$g'(0) = \frac{2}{1-t}\Big|_{t=0} = 2 = \mu_X \qquad g''(0) = \frac{2}{(1-t)^2}\Big|_{t=0} = 2 = \sigma_X^2$$

$$E[(X-\mu)^3] = g'''(0) = \frac{4}{(1-t)^3}\Big|_{t=0} = 4 \rightarrow \gamma_X = \frac{4}{(\sqrt{2})^3} = \sqrt{2}$$

(d) 3^{rd} quantile of X ?

Def $\rightarrow$ π_p 100^{th} percentile if $P(X \le \pi_p) = p$

$$\therefore \text{Find } \pi_{0.75} = P(X \le \pi_{0.75}) = 0.75 \rightarrow \int_0^{\pi_{0.75}} x e^{-x}\,dx = 0.75$$

Wednesday 2/21

$P(X \wedge Y \wedge Z)$ $\downarrow$ min $\qquad$ $P((X \wedge Y) \vee Z)$ $\downarrow$ max

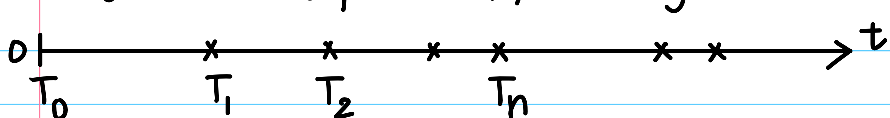
$X_1 \sim f_1(x) = e^{-x},\ x \ge 0 \qquad X_2 \sim f_2(x) = xe^{-x},\ x \ge 0 \qquad c_1 = 1/3,\ c_2 = 2/3$

$\therefore f(x) = \frac{1}{3}e^{-x} + \frac{2}{3}xe^{-x},\ x > 0 \longrightarrow$ convex combination of distributions

Monday 2/26

3.2 The Exponential, Gamma and Chi-square distributions

Consider the process w/ intensity λ



T_n = n^{th} arrival time of phone calls

(5) $T_n \sim \text{Gamma}(n, \lambda)$

$$\boxed{\begin{array}{l} \text{pdf: } f_{T_n}(n) = \dfrac{\lambda^n t^{n-1} e^{-\lambda t}}{\Gamma(n)} \\ \text{def: } \Gamma(\alpha) = \displaystyle\int_0^\infty x^{\alpha-1} e^{-x} dx \end{array}}$$

$t > 0$

$\alpha > 0$

(1) $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1), \ \alpha > 1$

Proof $\rightarrow$

$$\int_0^\infty x^{\alpha-1} e^{-x} dx = (-e^{-x}) x^{\alpha-1} \Big|_0^\infty - \int_0^\infty (-e^{-x})(\alpha-1) x^{\alpha-2} dx$$

$$= (\alpha-1)\int_0^\infty x^{\alpha-2} e^{-x} dx = (\alpha-1)\Gamma(\alpha-1)$$

(2) $\Gamma(1) = 1$

Proof $\rightarrow$

$$\int_0^\infty x^{1-1} e^{-x} dx = \int_0^\infty e^{-x} dx = 1$$

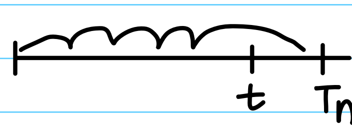
(3) $\Gamma(n) = (n-1)!$

Proof $\rightarrow$

$$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = \dots (n-1)(n-2)\dots 1 = (n-1)!$$

(4) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

e.g: $\Gamma(\underset{\uparrow \atop \alpha}{\frac{3}{2}}) = (\frac{3}{2}-1)\Gamma(\frac{3}{2}-1) = \frac{1}{2}\Gamma(\frac{1}{2}) = \frac{1}{2}\sqrt{\pi}$



Proof Gamma $\rightarrow$

Step 1: Find cdf of T_n

$$F_{T_n}(t) = P(T_n \le t) = 1 - P(T_n > t) = 1 - P(\underset{\downarrow \atop \approx \text{Poisson}(\lambda t)}{N_{[0,t]}} \le n-1)$$

$$= 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

Step 2: Find pdf of T_n

$$f_{T_n}(t) = \frac{d}{dt}F_{T_n}(t) = -\sum_{k=0}^{n-1}\frac{k}{k!}(\lambda t)^{k-1}\lambda e^{-\lambda t} - \sum_{k=0}^{n-1}\frac{(\lambda t)^k}{k!}(-\lambda)e^{-\lambda t}$$

$$= \sum_{k=0}^{n-1}\frac{t^k\lambda^{k+1}}{k!}e^{-\lambda t} - \sum_{k=1}^{n-1}\frac{t^{k-1}\lambda^k}{(k-1)!}e^{-\lambda t}$$

$$= \sum_{k=0}^{n-1}\frac{t^k\lambda^{k+1}}{k!}e^{-\lambda t} - \sum_{s=0}^{n-2}\frac{t^s\lambda^{k+1}}{s!}e^{-\lambda t} = \frac{t^{n-1}\lambda^n}{(n-1)!}e^{-\lambda t}$$

↳ $T_n \sim \text{Gamma}(n, \lambda)$ Q.E.D!

Wednesday 2/28

$X \sim \text{Gamma}(\alpha, \lambda)$ if pdf $f(x) = \dfrac{\lambda^\alpha x^{\alpha-1}e^{-\lambda x}}{\Gamma(\alpha)}$, $x \geq 0, \lambda > 0, \alpha > 0$

$\lambda = \frac{1}{\theta}$ or $\theta = \frac{1}{\lambda}$

↳ $X \sim \text{Gamma}(\alpha, \theta)$ if pdf $f(x) = \dfrac{x^{\alpha-1}}{\theta^\alpha\Gamma(\alpha)}e^{\frac{-x}{\theta}}$, $x \geq 0$

e.g.: $\text{Gamma}(\alpha, \lambda = 2) \longrightarrow \text{Gamma}(\alpha, \theta = \frac{1}{2})$

shape param — scale param

$\alpha = 1 \rightarrow X \sim f(x) = \boxed{\lambda e^{-\lambda x}} = \boxed{\frac{1}{\theta}e^{\frac{-x}{\theta}}}$, $x \geq 0$

$X \sim \text{Exp}(\lambda)$ $X \sim \text{Exp}(\theta)$

Properties → $X \sim \text{Gamma}(\alpha, \theta)$

(1) $M_X(t) = \dfrac{1}{(1-\theta t)^\alpha}$, $t < \frac{1}{\theta}$

Proof →

$$M_X(t) = E[e^{tX}] = \int_0^\infty e^{tx}\frac{x^{\alpha-1}}{\Gamma(\alpha)\theta^\alpha}e^{\frac{-x}{\theta}}dx = \frac{\theta'^\alpha}{\theta^\alpha}\int_0^\infty \frac{x^{\alpha-1}}{\Gamma(\alpha)\theta'^\alpha}e^{-\boxed{\left(\frac{1}{\theta}-t\right)}x}dx$$

$$\left(\frac{1}{\theta}-t\right) = \frac{1}{\theta'}$$

$$= \left(\frac{\theta'}{\theta}\right)^\alpha\int_0^\infty \text{Gamma}(\alpha, \theta')dx = \left(\frac{1}{\frac{1}{\theta}-t} \times \frac{1}{\theta}\right)^\alpha = \frac{1}{(1-\theta t)^\alpha},\ t < \frac{1}{\theta}$$

(2) $\mu_X = \alpha\theta$ (3) $\sigma_X^2 = \alpha\theta^2$ (4) $\gamma_X = \dfrac{2}{\sqrt{\alpha}}$

Proof $\rightarrow$

$$g(t) = \ln M_X(t) = \alpha \ln\left(\frac{1}{1-\theta t}\right) = -\alpha \ln(1-\theta t)$$

$$g'(t) = -\alpha \frac{-\theta}{1-\theta t} = \alpha\theta(1-\theta t)^{-1} \rightarrow \mu_X = g'(0) = \alpha\theta$$

$$g''(t) = \alpha\theta^2(1-\theta t)^{-2} \rightarrow \sigma_X^2 = g''(0) = \alpha\theta^2$$

$$g'''(t) = 2\alpha\theta^3(1-\theta t)^{-3} \rightarrow \gamma_X = \frac{g'''(0)}{\sigma_X{}^3} = \frac{2\alpha\theta^3}{(\alpha\theta^2)^{3/2}} = \frac{2}{\sqrt{\alpha}} > 0$$

Def $\rightarrow$ Chi-Square distribution $\boxed{X \sim \chi^2(r) = \text{Gamma}\left(\frac{r}{2}, \theta = 2\right)}$

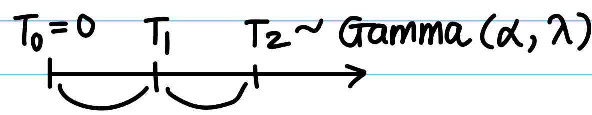
$$\mu_X = \alpha\theta = \frac{r}{2} \times 2 = r \qquad \gamma_X = \frac{2}{\sqrt{\alpha}} = 2 \times \sqrt{\frac{2}{r}} = \frac{2\sqrt{2}}{\sqrt{r}}$$

$$\sigma_X^2 = \alpha\theta^2 = \frac{r}{2} \times 2^2 = 2r \qquad M_X(t) = \frac{1}{(1-2t)^{r/2}}, \ t < \frac{1}{2}$$

$\boxed{3.2.4}$ Suppose #customers/hr arriving at a shop follows a Poisson process w/ mean 20. What is the prob. that the 2^{nd} customer arrives more than 5 mins after the shop opens for the day?

Gamma ①

$$\lambda = \frac{20}{hr} = \frac{20}{60 \text{ mins}} = \frac{1}{3 \text{ mins}}$$



$$P(T_2 > 5) = \int_5^\infty \frac{x^{2-1}\left(\frac{1}{3}\right)^2 e^{-1/3 x}}{\Gamma(2)} dx = \frac{1}{9}\int_5^\infty x e^{-1/3 x} dx$$

$\Gamma(n) = (n-1)! \rightarrow \Gamma(2) = 1! = 1$

$$= \frac{1}{9}\left(-3e^{-1/3 x} x \Big|_5^\infty - \int_5^\infty -3e^{-1/3 x} dx\right)$$

$$= \frac{1}{9}\left(15e^{-5/3} + 3(-3)e^{-1/3 x}\Big|_5^\infty\right) = \frac{1}{9}\left(15e^{-5/3} + 9e^{-5/3}\right) = \frac{8}{3}e^{-5/3}$$

same

Poisson ②

$$P(N_{[0,5]} \le 1) = P(N_{[0,5]} = 0) + P(N_{[0,5]} = 1) = \sum_{k=0}^{1} \frac{(5/3)^k}{k!} e^{-5/3}$$

$\uparrow$ Poisson(5. 1/3)

convert from Gamma to CS

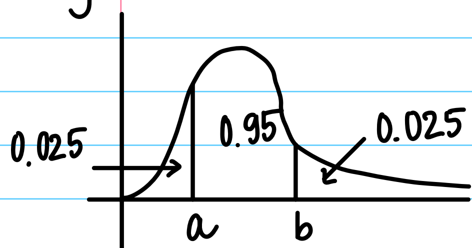
If $X \sim \text{Gamma}(\alpha, \lambda)$ then $Y = \frac{2X}{\theta} \sim \chi^2(2\alpha)$

<u>Proof</u> $\rightarrow$ $M_Y(t) = E[e^{tY}] = E[e^{\frac{t2X}{\theta}}] = E[e^{(\frac{2t}{\theta})X}] = \frac{1}{(1-\theta \cdot \frac{2t}{\theta})^{\alpha}} = \frac{1}{(1-2t)^{\alpha}}$

$= \frac{1}{(1-2t)^{2\alpha/2}} \sim \chi^2(2\alpha) \rightarrow \chi^2(r) = \frac{1}{(1-2t)^{r/2}}$

$2\alpha/2 \downarrow$ new r

<u>e.g</u>: $X \sim \chi^2(7)$. Find a, b such that $P(a < X < b) = 0.95$

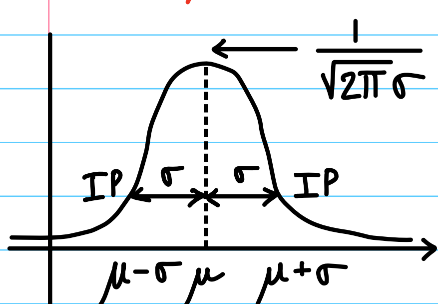


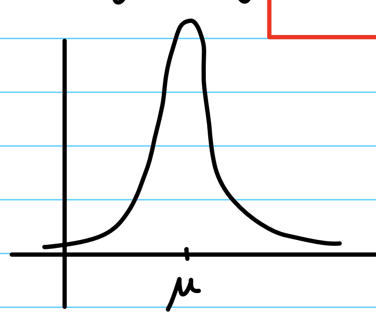
$\chi^2_{0.975} = 1.69 = a$ $\quad$ $\chi^2_{0.025} = 16.01 = b$

Monday 3/11

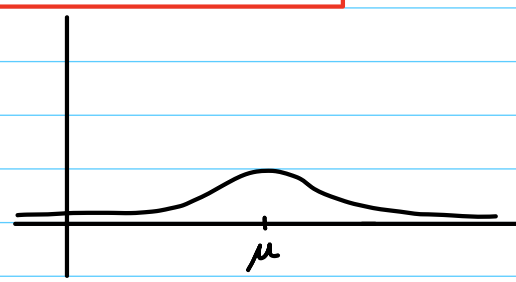
3.3 The Normal Distribution

<u>Def</u> $\rightarrow$ $X \sim N(\mu, \sigma^2)$ if the pdf is given by $\boxed{f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-1}{2\sigma^2}(x-\mu)^2}}$, $-\infty < x < \infty$





σ small



σ large

<u>Properties</u> $\rightarrow$ Let $X \sim N(\mu, \sigma^2)$

(a) $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$, $t \in \mathbb{R}$ $\qquad$ (c) $Var(X) = \sigma^2$

(b) $E[X] = \mu$ $\qquad$ (d) $\gamma = 0$

<u>Proof</u> $\rightarrow$ start w/ std dist.

$Z \sim N(0,1)$ is called a standard normal distribution

$\hookrightarrow$ Show that $M_Z(t) = e^{\frac{1}{2}t^2}$

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

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2 + \frac{1}{2}t^2} dx = e^{\frac{1}{2}t^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} dx}_{N(t,1)} = e^{\frac{1}{2}t^2}$$

Proof → (a) $M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$

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

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

$du = \frac{dx}{\sigma}$

½ times squared

$x = \mu + \sigma u$

$$= e^{t\mu} \underbrace{\int_{-\infty}^{\infty} e^{\boxed{t\sigma}u} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}u^2} du}_{N(0,1)} = e^{t\mu} e^{\frac{1}{2}t^2\sigma^2} = e^{t\mu + \frac{1}{2}\sigma^2 t^2}, \quad t \in \mathbb{R}$$

Q.E.D!

(b) Let $g(t) = \ln M_X(t) = t\mu + \frac{1}{2}\sigma^2 t^2$

$g'(t) = \mu + \sigma^2 t \longrightarrow E[X] = g'(0) = \mu$

$g''(t) = \sigma^2 \longrightarrow Var(X) = g''(0) = \sigma^2$

$g'''(t) = 0 \longrightarrow \gamma = \frac{g'''(0)}{\sigma^3} = 0$

Q.E.D!

Properties → Let $Z \sim N(0,1)$

(e) $E[Z^n] = \begin{cases} 0 & \text{if } n \text{ odd} \\ (2k-1)!! & \text{if } n = 2k \text{ even} \end{cases}$

$5!! = 5.3.1$

$7!! = 7.5.3.1$

Proof → $M_Z(t) = e^{\frac{1}{2}t^2} = \sum_{k=0}^{\infty} \frac{(\frac{1}{2}t^2)^k}{k!} = \sum_{k=0}^{\infty} \boxed{\frac{t^{2k}}{k!\,2^k}}$

also

$$= \sum_{n=0}^{\infty} \boxed{\frac{E[Z^n]}{n!}} t^n$$

compare

$\therefore E[Z^n] = 0$ if n odd

If $n = 2K$ even then $\frac{E[Z^{2K}]}{(2K)!} = \frac{1}{K!\,2^K}$

$$\rightarrow E[Z^{2K}] = \frac{(2K)!}{K!\,2^K} = \frac{\overbrace{(2K)(2K-1)(2K-2)(2K-3)\cdots 4.3.2.1}^{2^K K!}}{K!\,2^K} = \frac{2^K K!\,\overset{\text{remain odd terms}}{(2K-1)!!}}{K!\,2^K}$$

$= (2K-1)!!$ Q.E.D!

e.g: $\int_{-\infty}^{\infty} x^6 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 5!! = 5.3.1 = 15$ even $(n-1)!!$

$\int_{-\infty}^{\infty} x^5 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 0$ odd, symmetric shape

Theorem → $\boxed{\text{Let } X \sim N(\mu, \sigma^2) \text{ then } Z = \frac{X-\mu}{\sigma} \sim N(0,1)}$

$\boxed{\text{Let } Z \sim N(0,1) \text{ then } V = Z^2 \sim \chi^2(1)}$ $\chi^2(r) = \text{Gamma}(\frac{r}{2}, \theta = 2)$

Proof → (1) Find cdf of V

$F_V(v) = P(V \le v) = P(Z^2 \le v) = P(-\sqrt{v} \le Z \le \sqrt{v}) = F_Z(\sqrt{v}) - F_Z(-\sqrt{v})$

(2) Find pdf of V

$$f_V(v) = \frac{d}{dv} F_V(v) = \frac{d}{dv}\left(F_Z(\sqrt{v}) - F_Z(-\sqrt{v})\right)$$

$F_Z'(z) = f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v} \cdot \frac{1}{2\sqrt{v}} - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v} \cdot \frac{-1}{2\sqrt{v}}$$

$$= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{v}} e^{-\frac{1}{2}v}, \ v > 0 = \frac{v^{\frac{1}{2}-1}}{\underbrace{\Gamma(\frac{1}{2})}_{=\sqrt{\pi}} 2^{\frac{1}{2}}} e^{-v/2}, \ v > 0 \sim \chi^2(1)$$

Wednesday 3/13

4.1 Bivariate Distributions of the Discrete Type

Def → Let X, Y be 2 random variables defined in a discrete sample space. The joint pmf of X, Y is defined by

$\boxed{f(x,y) = P(X = x, Y = y) \text{ for } (x,y) \in S_{XY}}$

S_{XY} = space of (X, Y)
S_X = space of X

$$\boxed{P((X,Y)\in A) = \sum_{(x,y)\in A} f(x,y)}$$

e.g: Roll a pair of fair dice. Let X = smaller of the 2 dice, Y = larger of the 2 dice

(a) Find the joint pmf of X, Y

$f_Y(y)$	y \ x	1	2	3	4	5	6
$11/36$	6	$2/36$	$2/36$	$2/36$	$2/36$	$2/36$	$1/36$
$9/36$	5	$2/36$	$2/36$	$2/36$	$2/36$	$1/36$	
$7/36$	4	$2/36$	$2/36$	$2/36$	$1/36$		
$5/36$	3	$2/36$	$2/36$	$1/36$			
$3/36$	2	$2/36$	$1/36$				
$1/36$	1	$1/36$					$=0$
	$f_X(x)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

S_{XY}

restriction: Y always $\geq X$

(b) Find X-marginal pmf $= f_X(x) = \sum_y f(x,y)$ fixed X, sum Y

(c) Find Y-marginal pmf $= f_Y(y) = \sum_x f(x,y)$ fixed Y, sum X

(d) Find $P(X<Y) = 1 - P(X=Y) = 1 - \frac{6}{36} = \frac{30}{36}$

(e) Are X & Y independent? No

$$\boxed{\begin{array}{c} X, Y \text{ are independent if } f(x,y) = f_X(x)\cdot f_Y(x) \text{ for all } x,y \\ P(X=x \cap Y=y) = P(X=x)\cdot P(Y=y) \text{ for all } x,y \end{array}}$$

if independent $\longrightarrow S_{XY} = S_X \times S_Y$ (must be rectangular shape, not $\triangle$)

e.g: Suppose the joint pmf of X & Y is $f(x,y) = \frac{xy^2}{30}$, $x = 1,2,3$ $y = 1,2$

(a) Find $f_X(x) = \sum_y f(x,y) = \sum_{y=1}^{2} \frac{xy^2}{30} = \frac{x}{30} + \frac{4x}{30} = \frac{5x}{30}$, $x = 1,2,3$

(b) Find $f_Y(y) = \sum_x f(x,y) = \sum_{x=1}^{3} \frac{xy^2}{30} = \frac{y^2}{30} + \frac{2y^2}{30} + \frac{3y^2}{30} = \frac{6y^2}{30}$, $y = 1,2$

(c) Are X, Y inde ? Check $f_X(x) \cdot f_Y(y) = \frac{5x}{30} \cdot \frac{6y^2}{30} = \frac{xy^2}{30} = f(x,y) \to$ inde ✓

e.g: $f(x,y) = cxy^2$. Find c ?

$$\sum_{X=1}^{3}\sum_{Y=1}^{2} cxy^2 = c\sum_{X=1}^{3}(x \cdot 1^2 + x \cdot 2^2) = c\sum_{X=1}^{3} 5x = 5c(1+2+3) = 30c = 1$$

$\therefore c = \frac{1}{30}$

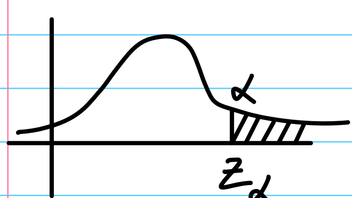
$\boxed{3.3.5}$ $X \sim N(6, 25)$. Find $P(6 \le X \le 14)$

$$Z = \frac{X-\mu}{\sigma} \sim N(0,1) \to P(6 \le X \le 14) = P\left(\frac{6-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{14-\mu}{\sigma}\right)$$

$$= P\left(\frac{6-6}{5} \le Z \le \frac{14-6}{5}\right) = P(0 \le Z \le 1.6) = P(Z \le 1.6) - P(Z \le 0)$$

$$= \Phi(1.6) - \Phi(0) = 0.952 - 0.5 = 0.452$$

Z_α upper quartile



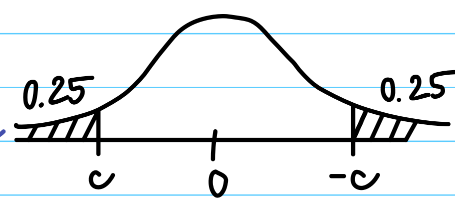
$P(Z \ge c) = 0.25$

$c = Z_{0.25} = 0.67$

the table gives the right hand side

$P(Z \le c) = 0.25$

$-c = Z_{0.25} = 0.67 \to c = -0.67$



$$P(X \le c) = P\left(\frac{X-\mu}{\sigma} \le \frac{c-\mu}{\sigma}\right) = P\left(Z \le \frac{c-6}{5}\right) = 0.25 \to \frac{c-6}{5} = -0.67$$

Monday 3/18

4.2 The Correlation Coefficient

Def →

Let X, Y be discrete random variables w/ joint pmf f(x,y) then

$$\boxed{E[u(x,y)] = \sum_{x,y} u(x,y) f(x,y)} \quad \boxed{\mu_X = E[X] = \sum_{x,y} x f(x,y) = \sum_{x,y} x f_X(x)}$$

Proof →

$$\mu_X = E[X] = \sum_{x,y} x f(x,y) = \sum_{x}\sum_{y} x f(x,y) = \sum_{x} x \sum_{y} y f(x,y) = \sum_{x} x f_X(x)$$

$$\mu_Y = E[Y] = \sum_{x,y} y f(x,y) = \sum_{y} y f_Y(y)$$

$$\sigma_X^2 = E[(X-\mu_X)^2] = E[X^2] - \mu_X^2 = \sum_{x} x^2 f_X(x) - \mu_X^2$$

$$\sigma_Y^2 = \sum_{y} y^2 f_Y(y) - \mu_Y^2$$

Def →

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

$$= \boxed{\sum_{x,y} xy f(x,y) - \mu_X\mu_Y}$$

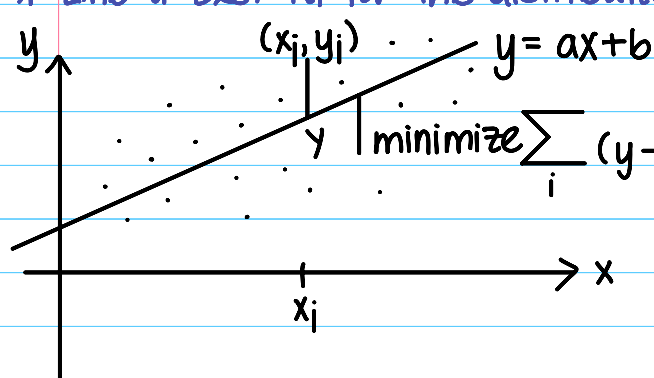
Correlation coefficient $\boxed{\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}}$ dimensionless

Properties →

(a) $Cov(X,Y) = E[XY] - \mu_X\mu_Y = \rho\sigma_X\sigma_Y$

(b) $E[XY] = \mu_X\mu_Y + \rho\sigma_X\sigma_Y$

* Line of best fit for the distribution $f(x,y)$



minimize $\sum_i (y-y_i)^2 f(x_i,y_i) = \sum_i [(ax_i+b)-y_i)^2] f(x_i,y_i)$

$$= E[(aX+b-Y)^2] \overset{set}{=} K(a,b)$$

* Find a, b such that $K(a,b)$ is a min

$$0 = \frac{\partial K(a,b)}{\partial a} = E[2(aX+b-Y)X] = E[(aX+b-Y)X]$$

$$0 = \frac{\partial K(a,b)}{\partial b} = E[2(aX+b-Y)] = E[(aX+b-Y)]$$

$\therefore 0 = aE[X^2] + b\mu_X - E[XY]$

$0 = a\mu_X + b - \mu_Y = a\mu_X^2 + b\mu_X - \mu_X\mu_Y$ subtract

μ_X

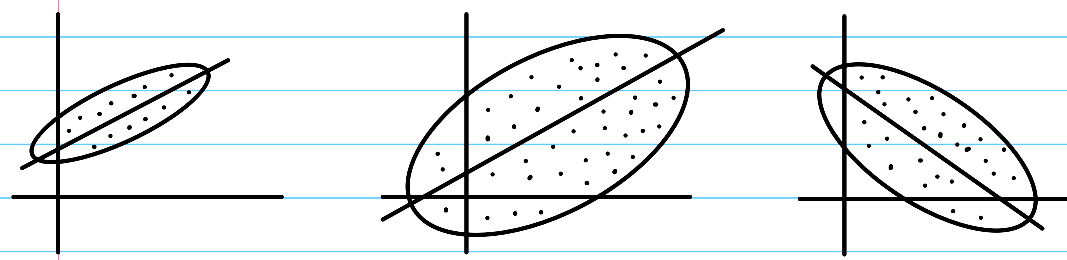
$\therefore\ 0 = a\sigma_X^2 - C_{XY} \longrightarrow a = \dfrac{C_{XY}}{\sigma_X^2}$

$\therefore\ 0 = \dfrac{C_{XY}}{\sigma_X^2}\mu_X + b - \mu_Y \rightarrow b = \mu_Y - \dfrac{C_{XY}}{\sigma_X^2}\mu_X$

$\therefore\ y = ax + b = \dfrac{C_{XY}}{\sigma_X^2}x + \mu_Y - \dfrac{C_{XY}}{\sigma_X^2}\mu_X = \dfrac{C_{XY}}{\sigma_X^2}(x-\mu_X) + \mu_Y = \rho\dfrac{\sigma_Y}{\sigma_X}(x-\mu_X) + \mu_Y$

is the line of best fit for dist. $f(x,y)$

* At the minimum: $K(a,b) = \sigma_Y^2(1-\rho^2)$



$\rho \sim 1, \rho > 0$ $\rho \sim 0, \rho > 0$ $\rho < 0$

Conclusion → If $\rho = \dfrac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\sigma_Y} = E\left[\left(\dfrac{X-\mu_X}{\sigma_X}\right)\left(\dfrac{Y-\mu_Y}{\sigma_Y}\right)\right] > 0$

then $\dfrac{X-\mu_X}{\sigma_X}$ & $\dfrac{Y-\mu_Y}{\sigma_Y}$ are more likely to have the same sign → positive correlation

* Calculate $K(a,b)$ at minimum

$$K(a,b) = E[(aX+b-Y)^2] = E\left\{\left[\rho\frac{\sigma_Y}{\sigma_X}(X-\mu_X) + \mu_Y - Y\right]^2\right\}$$

$$= \sigma_Y^2 E\left\{\left[\rho\left(\underbrace{\frac{X-\mu_X}{\sigma_X}}_{\tilde{X}}\right) - \left(\underbrace{\frac{Y-\mu_Y}{\sigma_Y}}_{\tilde{Y}}\right)\right]^2\right\}$$

normalize (imagine $N(0,1)$)

$$= \sigma_Y^2 E[(\rho\tilde{X} - \tilde{Y})^2]$$

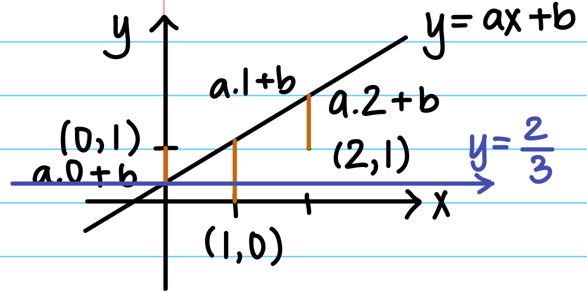
$$= \sigma_Y^2 E[\rho^2\tilde{X}^2 - 2\rho\tilde{X}\tilde{Y} + \tilde{Y}^2] = \sigma_Y^2(\rho^2 \cdot 1 - 2\rho\cdot\rho + 1) = \sigma_Y^2(1-\rho^2)$$

Properties → $E\left[\dfrac{X-\mu_X}{\sigma_X}\right] = 0$ $E\left[\left(\dfrac{X-\mu_X}{\sigma_X}\right)^2\right] = 1$ $E\left[\left(\dfrac{X-\mu_X}{\sigma_X}\right)\left(\dfrac{Y-\mu_Y}{\sigma_Y}\right)\right] = \rho$

→ formula

Wednesday 3/20 $\boxed{4.2.3}$ Let X, Y have the joint pmf $f(x,y) = \frac{1}{3}$ $(x,y) = (0,1), (1,0), (2,1)$

(a) Find $K(a,b)$



$$E[u(X,Y)] = \sum_{i=1}^{3} u(x_i, y_i) f(x_i, y_i)$$

$$\sum_i (y_i, \underset{\uparrow \atop ax_i+b}{y})^2 f(x_i, y_i) = \sum_i (y_i - (ax_i+b))^2 f(x_i, y_i)$$

$$= E[(Y - aX - b)^2]$$

$$K(a,b) = \frac{1}{3}(b-1)^2 + \frac{1}{3}(a+b-0)^2 + \frac{1}{3}(2a+b-1)^2$$

$$\left.\begin{aligned} 0 &= \frac{\partial K(a,b)}{\partial a} = \frac{2}{3}(a+b) + \frac{2}{3}(2)(2a+b-1) \longrightarrow 5a+3b-2=0 \\ 0 &= \frac{\partial K(a,b)}{\partial b} = \frac{2}{3}(b-1) + \frac{2}{3}(a+b) + \frac{2}{3}(2a+b-1) \longrightarrow 3a+3b-2=0 \end{aligned}\right\}$$

$\therefore a = 0,\ b = \frac{2}{3}$

(b) Find the line of best fit (regression line) $y = ax+b = \frac{2}{3}$

$$y = ax+b = \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) + \mu_Y = \frac{2}{3} \longrightarrow \rho = 0$$

$\not\Leftarrow$ does not imply backwards

Properties →

(a) If X, Y independent then $\rho = 0$
(b) $-1 \leq \rho \leq 1$

Proof → (a) $\rho = \frac{C_{XY}}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} = \frac{E[XY] - \mu_X \mu_Y}{\sigma_X \sigma_Y}$

$$E[XY] = \sum_{x,y} xy\, f(x,y) \overset{ind}{=} \sum_{x,y} xy\, f_X(x) f_Y(y) = \sum_x x f_X(x) \sum_y y f_Y(y) = \mu_X \mu_Y$$

$\therefore C_{XY} = 0 \quad \therefore \rho = 0$ Q.E.D!

(b) Step 1: Show that $|E[XY]| \leq \sqrt{E[X^2]}\sqrt{E[Y^2]}$ Cauchy-Schwarz

$$0 \leq E[(\alpha X + Y)^2] = E[\alpha^2 X^2 + 2\alpha XY + Y^2] = \alpha^2 E[X^2] + 2\alpha E[XY] + E[Y^2]$$

or $0 \le a\alpha^2 + b\alpha + c$ for all $\alpha \in \mathbb{R}$

$\hookrightarrow b^2 - 4ac = 4E[XY]^2 - 4E[X^2]E[Y^2] \le 0$) either 1 or no solution

$$E[XY]^2 \le E[X^2]E[Y^2]$$

$$|E[XY]| \le \sqrt{E[X^2]}\sqrt{E[Y^2]}$$

Step 2: Apply to $X-\mu_X$ & $Y-\mu_Y$

$$\therefore |E[(X-\mu_X)(Y-\mu_Y)]| \le \sqrt{E[(X-\mu_X)^2]}\sqrt{E[(Y-\mu_Y)^2]}$$

$$|Cov(X,Y)| \le \sigma_X \sigma_Y$$

$$\therefore |\rho| = \left|\frac{Cov(X,Y)}{\sigma_X\sigma_Y}\right| \le 1 \longrightarrow -1 \le \rho \le 1$$ Q.E.D!

Friday 3/22

4.3 Conditional Distributions

Def $\rightarrow$ Let $f(x,y)$ be the joint pmf of X and Y then the conditional pmf of X given that $Y=y$ is

$$\boxed{g(x\mid y) = \frac{f(x,y)}{f_Y(y)}} = \frac{P(X=x, Y=y)}{P(Y=y)} = P(X=x \mid Y=y)$$

fn of x y fixed

The conditional pmf of Y given that $X=x$ is

$$\boxed{h(y\mid x) = \frac{f(x,y)}{f_X(x)}} = \frac{P(Y=y, X=x)}{P(X=x)} = P(Y=y \mid X=x)$$

fn of y y fixed

(1) $P(X \in A \mid Y=y) = \sum_{x \in A} g(x\mid y)$ $E[u(X)\mid Y=y] = \sum_x u(x)g(x\mid y)$

(2) Conditional mean of X given $Y=y$ $\mu_{X\mid Y} = E[X\mid Y=y] = \sum_x x\,g(x\mid y)$

(3) Conditional variance of X given $Y=y$

$$\sigma^2_{X\mid Y} = \sum_x (x-\mu_{X\mid Y})^2 g(x\mid y) = \sum_x x^2 g(x\mid y) - \mu^2_{X\mid Y} = E[X^2\mid Y=y] - (E[X\mid Y=y])^2$$

$\boxed{4.3.1}$ Let pmf of X, Y be $f(x,y) = \frac{x+y}{21}$, $x = 1,2,3$ $y = 1,2$

(a) Find $g(x|y)$

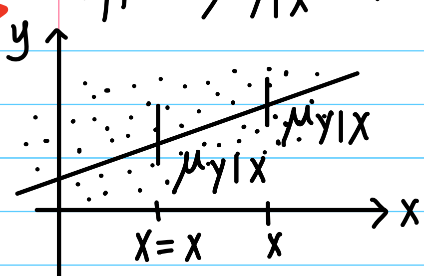
$$g(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{\frac{x+y}{21}}{\sum_{x=1}^{3} \frac{x+y}{21}} = \frac{x+y}{6+3y}, \quad x = 1,2,3 \quad y = 1,2$$

(b) Find $h(y|x)$

$$h(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{\frac{x+y}{21}}{\sum_{y=1}^{2} \frac{x+y}{21}} = \frac{x+y}{2x+3}, \quad x = 1,2,3 \quad y = 1,2$$

(c) $$\mu_{X|Y} = \sum x g(x|y) = \sum_{x=1}^{3} \frac{x(x+y)}{6+3y} = \frac{1(1+y)+2(2+y)+3(3+y)}{6+3y} = \frac{6y+14}{3y+6}, \quad y = 1,2$$

Theorem → Suppose $\mu_{Y|X} = ax+b$ then $\mu_{Y|X} = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X)$



Monday 3/25

$$\boxed{\begin{array}{l} (1)\ E[u(X,Y) \mid X = x] = \sum_y u(x,y)\, h(y|x) \\ (2)\ E[u(X,Y) \mid Y = y] = \sum_x u(x,y)\, g(x|y) \end{array}}$$

$$E[u(X,Y) \mid X = x] = \sum_y u(x,y)\, h(y|x) = \sum_y u(x,y) \frac{f(x,y)}{f_X(x)} \rightarrow \text{fn of } x$$

↘ known x

Define →

$$E[u(X,Y) \mid X] = \sum_y u(X,y) \frac{f(X,y)}{f_X(X)} = \sum_y u(X,y)\, h(y|X) \rightarrow \text{fn of } X$$

↓ a random variable dependent on X, unknown x

Theorem $\rightarrow$

$$\boxed{\begin{array}{l}(1)\ E\big[\underbrace{E[u(X,Y)\mid X]}_{\text{fn of } X}\big]=E[u(X,Y)] \text{ or } E\big[\underbrace{E[u(X,Y)\mid Y]}_{\text{fn of } Y}\big]=E[u(X,Y)] \\ \\ (2)\ E\big[u(X)v(Y)\mid X\big]=u(X)E\big[v(Y)\mid X\big]\end{array}}$$

Proof $\rightarrow$ (a) $E\big[E[u(X,Y)\mid X]\big]=E\big[\sum_y u(X,y)\,h(y\mid X)\big]$

$$=\sum_x\sum_y u(x,y)\,h(y\mid x)\,f_X(x)$$

$$=\sum_x\sum_y u(x,y)\,\frac{f(x,y)}{\cancel{f_X(x)}}\,\cancel{f_X(x)}=E[u(X,Y)] \quad \text{Q.E.D!}$$

(b) Subtask: Prove that $E\big[u(X)v(Y)\mid X\big]=u(X)E\big[v(Y)\mid X\big]$ on the set that $X=x$

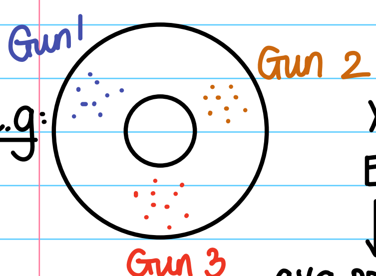
$\therefore$ On the set that $X=x$: LHS $=E\big[u(X)v(Y)\mid X\big]$

given x value, u(x) becomes a const

$=E\big[u(x)v(Y)\mid X=x\big]$ $\quad\therefore$ could factor out

$=u(x)E\big[v(Y)\mid X=x\big]=$ RHS at $X=x$ Q.E.D!

e.g:



$X=1,2,3$

$E[Y\mid X=1]=$ blue dots $\longrightarrow E\big[E[Y\mid X]\big]=$ avg of all dots

$=E[Y]$

$E[Y\mid X=1]$ $\downarrow$ avg points given specific gun

$E\big[E[Y\mid X]\big]$ $\downarrow$ any guns

w/ same prob.

Wednesday 3/27

$$\boxed{E[Y]=E\big[E[Y\mid X]\big]} \qquad \boxed{Var(Y)=E\big[Var(Y\mid X)\big]+Var\big(E[Y\mid X]\big)}$$

$$V \quad = \quad EV+VE$$

$\boxed{4.3.5}$ $X\sim$ Poisson(4). Suppose Y is a random variable where conditional prob. given that $X=x$ is Binomial$(x+1,p)$. Find $Var(Y)$

$Y\big|_{X=x}\sim Bin(x+1,p)\longrightarrow \mu_{Y\mid x}=(x+1)p\longrightarrow \mu_{Y\mid X}=(X+1)p$

$\sigma^2_{Y\mid x}=(x+1)pq\longrightarrow \sigma^2_{Y\mid X}=(X+1)pq$

$$Var(Y) = E[Var(Y|X)] + Var(E[Y|X])$$
$$= E[(X+1)pq] + Var((X+1)p)$$
$$= pqE[(X+1)] + Var(pX)$$
$$= pq(\lambda+1) \quad + p^2 Var(X) = pq(4+1) + p^2(4) = 5pq + 4p^2$$

Proof $\rightarrow$ $RHS = E[Var(Y|X)] + Var(E[Y|X])$

$$= E[E[Y^2|X] - \cancel{E[Y|X]^2}] + E[\cancel{E[Y|X]^2}] - E[E[Y|X]]^2$$
$$= E[Y^2] - E[Y]^2 = Var(Y) \quad \text{Q.E.D!}$$

Monday 4/1

4.4 Bivariate Distributions of Continuous Type

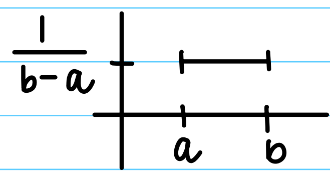
Let X, Y be continuous random variables w/ joint pdf

$f(x,y)$ $\quad (x,y) \in S \subseteq \mathbb{R}^2$ then $\quad$ S = space of (X,Y)

(a) $f(x,y) \geq 0$ $\quad f(x,y) = 0 \ \forall (x,y) \notin S$ $\quad$ = support of (X,Y)

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

(c) $P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$ for all $A \subseteq \mathbb{R}^2$

e.g: $U \sim Unif(a,b)$ pdf  $S = (a,b)$

X-marginal pdf: $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$ $\quad$ Y-marginal pdf: $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$

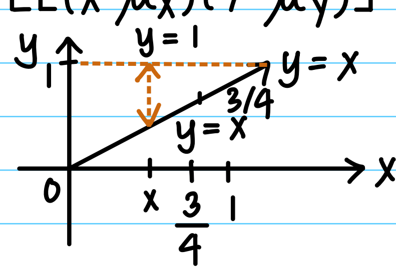
$$E[u(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x,y) f(x,y)\,dx\,dy$$

$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y)\,dx\,dy$ $\quad$ $Cov(X,Y) = E[XY] - \mu_X\mu_Y$ $\quad$ $\rho = \dfrac{Cov(X,Y)}{\sigma_X\sigma_Y}$

$$= E[(X-\mu_X)(Y-\mu_Y)]$$

4.4.6 Let $f(x,y) = 2$, $0 < x < y < 1$

(a) $f_X(x)$



$$f_X(x) = \int_x^1 f(x,y)\,dx = \int_x^1 2\,dx = 2x \Big|_x^1 = 2(1-x), \quad 0<x<1$$

(with $0<x<1$ written above)

(b) $$\mu_X = \int_0^1 x f_X(x)\,dx = \int_0^1 x\,2(1-x)\,dx = 2\int_0^1 (x-x^2)\,dx = 2\left(\frac{x^2}{2}-\frac{x^3}{3}\right)\Big|_0^1 = \frac{1}{3}$$

(c) $$\sigma_X^2 = E[X^2]-\mu_X^2 = \int_0^1 x^2 f_X(x)\,dx - \frac{1}{9} = \int_0^1 2(x^2-x^3)\,dx - \frac{1}{9} = \frac{1}{18}$$

(d) $$h(y\mid x) = \frac{f(x,y)}{f_X(x)} = \frac{2}{2(1-x)} = \frac{1}{1-x}, \quad x<y<1,\ 0<x<1$$

(e) $$E[Y\mid X=x] = \int_x^1 y\,h(y\mid x)\,dy = \int_x^1 \frac{y}{1-x}\,dy = \frac{1}{1-x}\left(\frac{y^2}{2}\right)\Big|_x^1 = \frac{1-x^2}{2(1-x)} = \frac{1+x}{2}$$
$= \mu_{Y|X}$

OR $Y|_X \sim \text{Unif}(x,1) \rightarrow \mu_{Y|X} = \dfrac{1+x}{2}$

(f) $$\text{Var}(Y\mid X=x) = \sigma^2_{Y|X} = \frac{(b-a)^2}{12} = \frac{(1-x)^2}{12}$$

(g) $$P\left(\frac{3}{4}<Y<\frac{7}{8}\;\Big|\;X=\frac{3}{4}\right) = \frac{7/8-3/4}{1-3/4} = \frac{1}{2}$$

get length then divide by total length of x
consider $Y\sim$ unif distribution

Def → $\boxed{\text{If } w(x) = E[u(X,Y)\mid X=x] \text{ then } w(X) = E[u(X,Y)\mid X]}$

4.5 The Bivariate Normal Distribution

Def → Random variables X, Y are said to have the bivariate normal distribution if the joint pdf

$$\boxed{f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\, e^{\frac{-1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}}}$$

$\mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y > 0$, $-1<\rho<1$ $\quad \tilde{x} \quad\quad -\infty < x, y < \infty \quad \tilde{y}$

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-1}{2(1-\rho^2)}\{\tilde{x}^2 - 2\rho\tilde{x}\tilde{y} + \tilde{y}^2\}}$$

Properties $\rightarrow$

$$\boxed{\begin{aligned} &(a)\ f_X(x) \sim N(\mu_X, \sigma_X^2) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{\frac{-1}{2\sigma_X^2}(x-\mu_X)^2},\quad x \in \mathbb{R} \\ &(b)\ f_Y(y) \sim N(\mu_Y, \sigma_Y^2) \\ &(c)\ f_{Y|X}(y|x) \sim N(\mu_{Y|X}, \sigma_{Y|X}^2) \rightarrow \mu_{Y|X} = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X) \\ &\qquad\qquad \sigma_{Y|X}^2 = \sigma_Y^2(1-\rho^2) \qquad \rho_{XY} = \rho \end{aligned}}$$

Wednesday 4/3

Proof $\rightarrow$

$$(a)\ f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-1}{2(1-\rho^2)}\{\tilde{x}^2 - 2\rho\tilde{x}\tilde{y} + \tilde{y}^2\}}$$

$$= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-1}{2}\cdot\frac{1}{1-\rho^2}\{(\tilde{y}-\rho\tilde{x})^2 - \rho^2\tilde{x}^2 + \tilde{x}^2\}}$$

$$= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-1}{2}\cdot\frac{1}{1-\rho^2}\left\{\left(\frac{y-\mu_Y}{\sigma_Y} - \rho\frac{x-\mu_X}{\sigma_X}\cdot\frac{\sigma_Y}{\sigma_Y}\right)^2 - (1-\rho^2)\tilde{x}^2\right\}}$$

$$= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-1}{2}\cdot\frac{1}{1-\rho^2}\cdot\frac{1}{\sigma_Y^2}\{\underbrace{y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)}_{m\ (\text{regression line})}\}^2 - \frac{-1}{2}\cdot\frac{(1-\rho^2)\tilde{x}^2}{1-\rho^2}}$$

$$= \frac{1}{2\pi\sigma_X} e^{\frac{-1}{2}\tilde{x}^2} \underbrace{\frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-1}{2}\cdot\frac{1}{1-\rho^2}\cdot\frac{1}{\sigma_Y^2}(y-m)^2}}_{Y\sim N(m,\ \sigma_X^2(1-\rho^2))}$$

$$= \frac{1}{2\pi\sigma_X} e^{\frac{-1}{2}\tilde{x}^2}$$

$$\therefore f_X(x) = \int_y f(x,y)\,dy = \frac{1}{2\pi\sigma_X} e^{\frac{-1}{2}\tilde{x}^2} = \frac{1}{2\pi\sigma_X} e^{\frac{-1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} \sim N(\mu_X, \sigma_X^2)$$

Q.E.D!

(c) $f_{Y|X}(y|x) = \dfrac{f(x,y)}{f_X(x)} \sim N(m, \sigma_Y^2(1-\rho^2))$, $m = \mu_Y + \rho\dfrac{\sigma_Y}{\sigma_X}(x-\mu_X)$

$$E[XY] = E\big[E[XY|X]\big] = E\big[X\,E[Y|X]\big]$$

$$= E\left\{X\left[\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(X-\mu_X)\right]\right\}$$

$$= \mu_Y\mu_X + \rho\frac{\sigma_Y}{\sigma_X}(\sigma_X^2 + \cancel{\mu_X^2} - \cancel{\mu_X^2}) = \mu_Y\mu_X + \rho\sigma_Y\sigma_X$$

$\therefore C_{X,Y} = E[XY] - \mu_X\mu_Y = \rho\sigma_Y\sigma_X \quad \therefore \rho_{X,Y} = \dfrac{C_{X,Y}}{\sigma_X\sigma_Y} = \dfrac{\rho\sigma_X\sigma_Y}{\sigma_X\sigma_Y} = \rho$ Q.E.D!

Theorem $\rightarrow$

Let X, Y have a bivariate normal distribution then
X, Y are independent $\leftrightarrow \rho = 0$

Def $\rightarrow$

Let X, Y be any random variables
X, Y are uncorrelated if $\rho_{XY} = 0$

X, Y independent $\longrightarrow$ X, Y uncorrelated
$\not\leftarrow$ not other way

$\boxed{4.5.2}$ X = high school grade, Y = 1st year college grade. Assume that X, Y have a bivariate normal distribution w/ $\mu_X = 2.9$, $\mu_Y = 2.4$, $\sigma_X = 0.4$, $\sigma_Y = 0.5$, $\rho = 0.6$

(a) Find $P(2.1 < Y < 3.3)$

$$= P\left(\frac{2.1-\mu_Y}{\sigma_Y} < \frac{Y-\mu_Y}{\sigma_Y} < \frac{3.3-\mu_Y}{\sigma_Y}\right) = P\left(\frac{2.1-2.4}{0.5} < Z < \frac{3.3-2.4}{0.5}\right) = 0.6898$$

(b) $P(2.1 < Y < 3.3 \mid X = 3.2)$

$\mu_Y + \rho\dfrac{\sigma_Y}{\sigma_X}(x-\mu_X)$

$$= P\left(\frac{2.1-\mu_{Y|X=3.2}}{\sigma_{Y|X=3.2}} < \frac{Y-\mu_{Y|X=3.2}}{\sigma_{Y|X=3.2}} < \frac{3.3-\mu_{Y|X=3.2}}{\sigma_{Y|X=3.2}}\right)$$

$\sqrt{\sigma_Y^2(1-\rho^2)}$

$$= P\left(\frac{2.1-(2.4+0.6\frac{0.5}{0.4}(3.2-2.9))}{\sqrt{0.5^2(1-0.6^2)}} < Z < \frac{3.3-(2.4+0.6\frac{0.5}{0.4}(3.2-2.9))}{\sqrt{0.5^2(1-0.6^2)}}\right)$$

$= 0.8594$

Friday 4/5

$$P(A) = \underbrace{E[\overbrace{P(A|X)}^{\text{fn of } X}]}_{\text{indication fn of an event}} = \int_{-\infty}^{\infty} P(A|X=x) \cdot f_X(x)\,dx$$

Event $A \subseteq S$ sample space $\longrightarrow$ Define $\chi_A(w) = \begin{cases} 1 & \text{if } w \in A \\ 0 & \text{if } w \notin A \end{cases}$ is called the indication fn

$\chi_A : S \to \mathbb{R}$ $\quad \therefore \chi_A$ is a random variable

$E[\chi_A] = P(A) = 1 \cdot P(\chi_A = 1) + 0 \cdot P(\chi_A = 0) = 1 \cdot P(A) + 0 \cdot P(A^c) = P(A)$

Theorem $\rightarrow$

rectangle set

$$\boxed{f(x,y) = h(x) \cdot g(y) \ \& \ S = S_X \times S_Y \longleftrightarrow X, Y \text{ independent}}$$

could rewrite as just x & just y

Monday 4/8

5.1 Functions of One Random Variables

* CDF technique

e.g.: $X \sim \text{Gamma}(\alpha, \theta)$. $Y = e^X$. Find pdf of Y

(1) find cdf of Y

$F_Y(y) = P(Y \leq y) = P(e^X \leq y) = P(X \leq \ln y) = F_X(\ln y)$

(2) find pdf of Y

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(\ln y) = f_X(\ln y) \cdot \frac{d(\ln y)}{dy} = \frac{(\ln y)^{\alpha-1}}{\Gamma(\alpha)\theta^\alpha} e^{-\ln y/\theta} \frac{1}{y}, \quad 1 < y < \infty$$

$$= \frac{(\ln y)^{\alpha-1}}{\Gamma(\alpha)\theta^\alpha} y^{-1/\theta} \left(\frac{1}{y}\right) = \frac{(\ln y)^{\alpha-1}}{\Gamma(\alpha)\theta^\alpha} \cdot \frac{1}{y^{1+1/\theta}}, \quad 1 < y < \infty$$

$\downarrow$ log-gamma distribution

* Change-of-variable technique

Suppose $y = g(x) \rightarrow x = g^{-1}(y)$

Given $X \sim f_X(x)$. Let $Y = g(X)$ then $\boxed{f_Y(y) = \sum_{x:\, g(x) = y} f_X(x) \frac{1}{|g'(x)|} \Bigg|_{x = g^{-1}(y)}}$

$y = e^x = g(x) \longrightarrow x = \ln y$

$$f_Y(y) = f_X(x) \cdot \frac{1}{e^x}\Big|_{x=\ln y} = \frac{(\ln y)^{\alpha-1}}{\Gamma(\alpha)\theta^\alpha} e^{-\ln y/\theta} \frac{1}{e^{\ln y}} = \frac{(\ln y)^{\alpha-1}}{\Gamma(\alpha)\theta^\alpha} \frac{1}{y^{1+1/\theta}}$$

5.2 $X = (X_1, X_2) \sim$ joint pdf $f(x_1, x_2)$

$Y = (Y_1, Y_2) \sim$ joint pdf $f(y_1, y_2)$

$$\begin{cases} x_1 = x_1(y_1, y_2) \\ x_2 = x_2(y_1, y_2) \end{cases} \qquad \begin{cases} y_1 = y_1(x_1, x_2) \\ y_2 = y_2(x_1, x_2) \end{cases}$$

$$\therefore f_Y(y_1, y_2) = f_X(x_1, x_2)\left|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right| = f_X(x_1, x_2)\frac{1}{\left|\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)}\right|}$$

where $\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \det\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = J$ & $\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = \det\begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{vmatrix}$

e.g: $f(x_1, x_2) = 2$. Let $Y_1 = \frac{X_1}{X_2}$, $Y_2 = X_2$ $(0 < x_1 < x_2 < 1)$

(a) joint pdf of Y_1, Y_2

Step 1 Find the joint range of (Y_1, Y_2)

$0 < y_1 < 1$ $\longrightarrow$ test: $Y_1 = \frac{X_1}{Y_2} \longrightarrow Y_1 Y_2 = X_1 \longrightarrow \underbrace{0 < y_1 y_2 < y_2 < 1}$

$0 < y_2 < 1$

satisfied

$$f_Y(y_1, y_2) = f_X(x_1, x_2)\left|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right| = 2\left|\det\begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix}\right| = 2|y_2| = 2y_2, \quad 0 < y_1 < 1,\ 0 < y_2 < 1$$

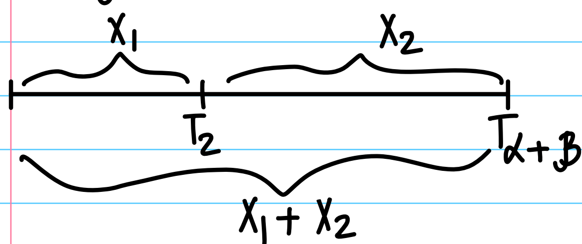
(b) $f_{Y_2}(y_2) = \int_0^1 f(y_1, y_2)\,dy_1 = \int_0^1 2y_2\,dy_1 = 2y_2, \quad 0 < y_2 < 1$

(c) $f_{Y_2}(y_1) = \int_0^1 f(y_1, y_2)\,dy_2 = \int_0^1 2y_2\,dy_2 = y_2^2\Big|_0^1 = 1, \quad 0 < y_1 < 1$

$\boxed{5.2.3}$ $X_1 \sim \text{Gamma}(\alpha, \theta)$ $\quad$ X_1, X_2 independent. Let $Y_1 = \dfrac{X_1}{X_1+X_2} \sim \text{Beta}(\alpha, \beta)$ $\quad$ $\nearrow > 0$

$X_2 \sim \text{Gamma}(\beta, \theta)$

(a) Find joint pdf of Y_1, Y_2 $\qquad$ $Y_2 = X_1 + X_2 \sim \text{Gamma}(\alpha+\beta, \theta)$ $\quad \rightarrow > 0$



$\dfrac{T_3}{T_{10}} \sim \text{Beta}(3, 7)$

Wednesday 4/10

$Y_1 = \dfrac{X_1}{Y_2} \rightarrow X_1 = Y_1 Y_2$

$X_2 = Y_2 - X_1 = Y_2 - Y_1 Y_2 = Y_2(1 - Y_1)$

$0 < x_1 < \infty \qquad 0 < x_2 < \infty$

$0 < y_1 y_2 < \infty \qquad 0 < \underbrace{y_2}_{\text{already} > 0}(1 - y_1) < \infty \rightarrow 0 < 1 - y_1 \rightarrow 0 < y_1 < 1$

$$f_Y(y_1, y_2) = f_X(x_1, x_2) \left| \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right|$$

X_1, X_2 inde $\therefore f_X(x_1) . f_X(x_2)$ $\leftarrow$

$$= \frac{x_1^{\alpha-1} x_2^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)\theta^\alpha \theta^\beta} e^{-\left(\frac{x_1}{\theta} + \frac{x_2}{\theta}\right)} \begin{vmatrix} y_2 & y_1 \\ -y_2 & 1-y_1 \end{vmatrix}$$

$$= \frac{(y_1 y_2)^{\alpha-1} (y_2(1-y_1))^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)\theta^\alpha \theta^\beta} e^{-y_2/\theta} |y_2(1-y_1) + y_1 y_2|$$

$$= \frac{y_1^{\alpha-1} y_2^{\alpha-1} y_2^{\beta-1} (1-y_1)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)\theta^\alpha \theta^\beta} e^{-y_2/\theta} \, y_2$$

since > 0

$$= \underbrace{\frac{y_1^{\alpha-1}(1-y_1)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}}_{} \; \underbrace{\frac{y_2^{\alpha+\beta-1}}{\theta^{\alpha+\beta}} e^{-y_2/\theta}}_{}, \quad 0 < y_1 < 1, \; 0 < y_2 < \infty$$

could separate them $\rightarrow Y_1, Y_2$ inde

$$= \frac{\Gamma(\alpha+\beta) y_1^{\alpha-1}(1-y_1)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \; \frac{y_2^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)\theta^{\alpha+\beta}} e^{-y_2/\theta}$$

add to represent Gamma$(\alpha+\beta, \theta)$

(b) $f_{Y_1}(y_1) = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y_1^{\alpha-1}(1-y_1)^{\beta-1}, \quad 0 < y_1 < 1$

$\sim \text{Beta}(\alpha, \beta)$

$W \sim F(r_1, r_2)$

5.2.4 Let $U \sim \chi^2(r_1)$, $V \sim \chi^2(r_2)$, U, V inde. Let $W = \dfrac{U/r_1}{V/r_2}$

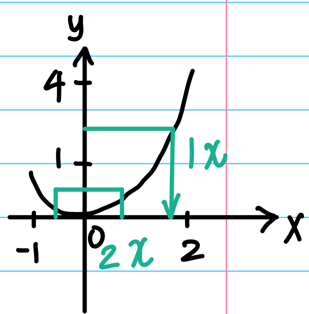
will be given
no need to
memorize

$$\boxed{f_W(w) = \frac{\Gamma\left(\frac{r_1+r_2}{2}\right)\left(\frac{r_1}{r_2}\right)^{r_2/2} w^{r_1/2 - 1}}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)\left(1+\frac{r_1}{r_2}w\right)^{\frac{r_1+r_2}{2}}}, \quad w > 0}$$ F distribution

5.1.8 Let $X \sim f(x) = \dfrac{x^2}{3}$, $-1 < x < 2$. Let $Y = X^2$. Find pdf of Y

$y = g(x) = x^2 \longrightarrow x = g^{-1}(y) = \pm\sqrt{y}$, $g'(x) = 2x$

$$f_Y(y) = \sum_y f_X(x) \cdot \frac{1}{|g'(x)|}\Bigg|_{x=g^{-1}(y)} = \frac{x^2}{|6x|}\Bigg|_{x=\sqrt{y}} + \underbrace{\frac{x^2}{|6x|}\Bigg|_{x=-\sqrt{y}}}_{=0 \text{ for } 2^{nd} \text{ condition}}$$



$$= \begin{cases} 0 & , y > 4 \\ \dfrac{(\sqrt{y})^2}{|6\sqrt{y}|} & , 1 < y < 4 \\ \dfrac{(\sqrt{y})^2}{|6\sqrt{y}|} + \dfrac{(-\sqrt{y})^2}{|-6\sqrt{y}|} & , 0 < y < 1 \end{cases} = \begin{cases} 0 & , y > 4 \\ \frac{1}{6}\sqrt{y} & , 1 < y < 4 \\ \frac{1}{3}\sqrt{y} & , 0 < y < 1 \end{cases}$$

Monday
4/15

5.3 Several Independent Random Variables

Let $X_1, X_2 \ldots X_n$ be random variables defined on the same prob. space

Discrete $\longrightarrow$ Joint pmf $f(x_1, x_2 \ldots x_n) = P(X_1 = x_1, X_2 = x_2 \ldots X_n = x_n)$

Continuous $\longrightarrow$ Joint pdf

$$f(x_1, x_2 \ldots x_n) \geq 0 \ni \iint_A \ldots \int f(x_1, x_2 \ldots x_n)\, dx_1 \ldots dx_n = P((X_1 \ldots X_n) \in A)$$
for all $A \subseteq \mathbb{R}^n$

$$E[u(X_1, \ldots X_n)] = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} u(x_1 \ldots x_n) f(x_1 \ldots x_n) dx_1 \ldots dx_n$$

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1 \ldots x_n) dx_2 \ldots dx_n \rightarrow f_{X_2 X_3}(x_2, x_3) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1 \ldots x_n) \underbrace{dx_1\, dx_4 \ldots}_{\text{skipping } x_2, x_3}$$

Def → $X_1, X_2 \dots X_n$ are independent if for all $A_i \subseteq \mathbb{R}$

$$P(X_1 \in A_1, X_2 \in A_2, \dots X_n \in A_n) = P(X_1 \in A_1) \dots P(X_n \in A_n)$$

$X_1, X_2 \dots X_n$ are independent $\leftrightarrow f(x_1, \dots x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$
$(X = (X_1 \dots X_n))$ and $S_X = S_{X_1} \times S_{X_2} \times \dots \times S_{X_n}$

Let X be a random variable w/ pdf or pmf $f(x)$
$(X_1, X_2 \dots X_n)$ is called a random sample of X of size n if $X_1, \dots X_n$ are inde. and $X_i \sim f(x)$ for all $i = 1, 2 \dots n$

$X_1, X_2 \dots X_n$ are independent & identically distributed random variables w/ common distribution $f(x)$

e.g: $X \sim U(0,1) \rightarrow$ Random samples: $X_1 \rightarrow x_1^{(1)}, x_1^{(2)} \dots x_1^{(m)}$
$X_2 \rightarrow x_2^{(1)}, x_2^{(2)} \dots x_2^{(m)}$
$\dots$
$X_n \rightarrow x_n^{(1)}, x_n^{(2)} \dots x_n^{(m)}$

Theorem → Let $X_1, X_2 \dots X_n$ be n random variables w/ means $\mu_1, \mu_2 \dots \mu_n$ and variances $\sigma_1^2, \sigma_2^2 \dots \sigma_n^2$

Let $Y = \sum_{i=1}^{n} a_i X_i$ then

(a) $\mu_Y = \sum_{i=1}^{n} a_i \mu_i$ (b) σ_Y^2

(c) If $X_1 \dots X_n$ are inde, $\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2$

Proof →

(a) $\mu_Y = E[Y] = E[\sum_{i=1}^{n} a_i X_i] = \sum_{i=1}^{n} a_i E[X_i] = \sum_{i=1}^{n} a_i \mu_i$ Q.E.D!

(b) $\sigma_Y^2 = E[(Y - \mu_Y)^2] = E[(\sum_{i=1}^{n} a_i X_i - \sum_{i=1}^{n} a_i \mu_i)^2]$

$= E[(\sum_{i=1}^{n} a_i (X_i - \mu_i))^2] = \sum_{i,j=1}^{n} a_i a_j E[(X_i - \mu_i)(X_j - \mu_j)] = \sum_{i,j=1}^{n} a_i a_j C_{ij}$

$= \underbrace{\sum_{i=1}^{n} a_i^2 \sigma_i^2}_{\text{same as } E[(X_i-\mu_i)^2] = \sigma_i^2} + \sum_{i \neq j} a_i a_j C_{ij} \xrightarrow[C_{ij} = 0]{\text{if } i,j \text{ inde}} \sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2$ Q.E.D!

Wednesday 4/17

Let $(X_1, X_2 \dots X_n)$ be a random sample of size n of X

- sample mean $\quad \bar{X} = \dfrac{X_1+\dots+X_n}{n}$
- sample variance $\quad s^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(x_i-\bar{X})^2$

Def $\rightarrow$ A fn $u(X_1, X_2, \dots X_n)$ is called a statistic

- $\bar{X}$ statistic for $E[X]$
- $\bar{X}$ estimator for $E[X]$
- s^2 estimator for σ_X^2

$$\boxed{\text{If } E[\bar{X}] = E[X] \text{ then } \bar{X} \text{ is an unbias estimator}}$$

Properties $\rightarrow$

$$\boxed{(a)\ \mu_{\bar{X}} = \mu_X \qquad (b)\ \sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \qquad (c)\ E[s^2] = \sigma_X^2}$$

Proof $\rightarrow$

(a) $\mu_{\bar{X}} = E[\bar{X}] = E\left[\dfrac{X_1+\dots+X_n}{n}\right] = \dfrac{1}{n}\Big(E[X_1]+\dots+E[X_n]\Big) = \dfrac{\mu_{X_1}+\dots+\mu_{X_n}}{n} = \mu_X$

(b) $\sigma_{\bar{X}}^2 = E[(\bar{X}-\mu_{\bar{X}})^2] = E\left[\left(\dfrac{X_1+\dots+X_n}{n} - \mu_X\right)^2\right]$

(a)

$$= E\left[\left(\frac{(X_1-\mu_X)+\dots+(X_n-\mu_X)}{n}\right)^2\right] = \frac{1}{n^2}\sum_{i,j=1}^{n} E[(X_i-\mu_X)(X_j-\mu_X)]$$

$$= \frac{1}{n^2}\sum_{i,j=1}^{n} E[(X_i-\mu_X)^2] = \frac{1}{n^2}(n\sigma_X^2) = \frac{\sigma_X^2}{n}$$

(c) Show that $E[s^2] = E\left[\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(x_i-\bar{X})^2\right] = \sigma_X^2$

$$n\sigma_X^2 = E\left[\sum_{i=1}^{n}(x_i-\mu_X)^2\right] = E\left[\sum_{i=1}^{n}(x_i-\bar{X}+\bar{X}-\mu_X)^2\right]$$

$$= E\left[\sum_{i=1}^{n}(x_i-\bar{X})^2 + 2\sum_{i=1}^{n}(x_i-\bar{X})(\bar{X}-\mu_X) + \sum_{i=1}^{n}(\bar{X}-\mu_X)^2\right]$$

$$= E\left[\sum_{i=1}^{n}(x_i-\bar{X})^2 + 2(\bar{X}-\mu_X)\sum_{i=1}^{n}(x_i-\bar{X}) + \sum_{i=1}^{n}(\bar{X}-\mu_X)^2\right]$$

$$= E\left[\sum_{i=1}^{n}(x_i-\bar{X})^2 + 2(\bar{X}-\mu_X)\left[\left(\sum_{i=1}^{n}x_i\right) - n\bar{X}\right]\right] + nE[(\bar{X}-\mu_X)^2]$$

$$= E\left[\sum_{i=1}^{n}(x_i-\bar{X})^2 + 2(\bar{X}-\mu_X)(x_1+x_2+\dots+x_n - n\frac{x_1+\dots+x_n}{n}\right] + n\frac{\sigma_X^2}{n}$$

$$\underbrace{}_{0}$$

$$= E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] + \sigma_{\bar{X}}^2 \longrightarrow E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = n\sigma_X^2 - \sigma_X^2 = (n-1)\sigma_X^2$$

$$\therefore E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \frac{\cancel{(n-1)}\sigma_X^2}{\cancel{n-1}} = \sigma_X^2 = E[S^2]$$ Q.E.D!

Friday 4/19

5.4 The Moment Generating Function Technique

Let $X_1, X_2 \dots X_n$ be independent random variables w/ respective m.g.f $M_{X_i}(t)$ $i = 1, 2 \dots n$

Theorem

$$\boxed{\text{Let } Y = \sum_{i=1}^{n} a_i X_i \text{ then } M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t)}$$

e.g: $X_1 \sim Gamma(\alpha, \theta)$ $\quad$ X_1, X_2 inde.

$X_2 \sim Gamma(\beta, \theta)$ $\quad$ Let $Y = X_1 + X_2$. Find distribution of Y

$$M_Y(t) = E[e^{tY}] = E[e^{t(X_1+X_2)}] = E[e^{tX_1} e^{tX_2}] \overset{\text{inde.}}{=} E[e^{tX_1}]E[e^{tX_2}]$$

$$= \frac{1}{(1-\theta t)^{\alpha}} \times \frac{1}{(1-\theta t)^{\beta}} = \frac{1}{(1-\theta t)^{\alpha+\beta}} \longrightarrow Y \sim Gamma(\alpha+\beta, \theta)$$

Proof

$$M_Y(t) = E[e^{tY}] = E[e^{t\sum_{i=1}^{n} a_i X_i}] \overset{\text{inde.}}{=} E\left[\prod_{i=1}^{n} e^{t a_i X_i}\right] = \prod_{i=1}^{n} E[e^{(a_i t) X_i}]$$

$$= \prod_{i=1}^{n} M_{X_i}(a_i t)$$ Q.E.D!

Properties

Let $X_1, X_2 \dots X_n$ be observations of a random sample from a distribution w/ m.g.f $M_X(t)$

Let $Y = X_1 + X_2 + \dots + X_n$ and $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$ then

(a) $M_Y(t) = [M_X(t)]^n$ $\quad$ (b) $M_{\bar{X}}(t) = [M_X(\frac{t}{n})]^n$

Theorem

Let $X_1, X_2 \dots X_n$ be inde. $X_i \sim \chi^2(r_i)$. Let $Y = X_1 + X_2 + \dots + X_n$ then $Y \sim \chi^2(r_1 + r_2 + \dots + r_n)$

same as Gamma

Q.E.D!

Proof $\rightarrow$ $M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} \frac{1}{(1-2t)^{r_i/2}} = \frac{1}{(1-2t)^{\sum r_i/2}} \longrightarrow Y \sim \chi^2(\sum_{i=1}^{n} r_i)$

Theorem $\rightarrow$ Let $Z_1, Z_2 \ldots Z_n$ be inde. standard normal distribution.
$W = Z_1^2 + Z_2^2 + \ldots + Z_n^2$ then $W \sim \chi^2(n)$

Theorem $\rightarrow$ Let $X_1, X_2 \ldots X_n$ be inde. and $X_i \sim N(\mu_i, \sigma_i^2)$ then

$$W = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \sim \chi^2(n)$$

$Z^2 \sim \chi^2(1) \xrightarrow{\text{sum } n} \chi^2(n)$

Theorem $\rightarrow$ $X_1, X_2 \ldots X_n$ be inde. and $X_i \sim Ber(p)$ then $Y = X_1 + X_2 + \ldots + X_n \sim Bin(n, p)$

Proof $\rightarrow$ $M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} (q + pe^t) = (q + pe^t)^n \longrightarrow Y \sim Bin(n, p)$

e.g: pmf of X: $f_X(x) = \frac{x}{6}$, $x = 1, 2, 3$ pmf of Y: $f_Y(y) = \begin{cases} 2/3, & y = 1 \\ 1/3, & y = 2 \end{cases}$

X, Y inde. Let $W = X + Y$. Find $f_W(w)$

$M_W(t) = M_X(t) M_Y(t) = \left(e^t \times \frac{1}{6} + e^{2t} \times \frac{2}{6} + e^{3t} \times \frac{3}{6}\right)\left(e^t \times \frac{2}{3} + e^{2t} \times \frac{1}{3}\right)$

$= e^{2t}\left(\frac{1}{6}\right)\left(\frac{2}{3}\right) + e^{3t}\left(\frac{1}{6} \times \frac{1}{3} + \frac{2}{6} \times \frac{2}{3}\right) + e^{4t}\left(\frac{2}{6} \times \frac{1}{3} + \frac{3}{6} \times \frac{2}{3}\right) + e^{5t}\left(\frac{3}{6}\right)\left(\frac{1}{3}\right)$

$\therefore f_W(2) = \frac{1}{6} \times \frac{2}{3}$

$f_W(3) = \frac{1}{6} \times \frac{1}{3} + \frac{2}{6} \times \frac{2}{3} \ldots$

Monday 4/22

5.5 Random Functions Associated w/ Normal Distributions

Let $X_1, \ldots X_n$ be independent, $X_i \sim N(\mu_i, \sigma_i^2)$

Let $Y = \sum_{i=1}^{n} c_i X_i$ then $Y \sim N\left(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2\right)$

Proof $\rightarrow$ $M_Y(t) = E[e^{tY}] = E[e^{t\sum_{i=1}^{n} c_i X_i}] = E[\prod_{i=1}^{n} e^{c_i t X_i}] \overset{\text{inde}}{=} \prod_{i=1}^{n} E[e^{c_i t X_i}]$

$$= \prod_{i=1}^{n} M_{X_i}(c_i t) = \prod_{i=1}^{n} e^{c_i \mu_i t + \frac{1}{2} c_i^2 \sigma_i^2 t^2}$$

$$= e^{(\sum_{i=1}^{n} c_i x_i) t + \frac{1}{2}(\sum_{i=1}^{n} c_i^2 \sigma_i^2) t^2} \rightarrow Y \sim N\left(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2\right)$$

Q.E.D!

Theorem $\rightarrow$ Let $X_1, X_2 \dots X_n$ be a random sample of size n from $N(\mu_i, \sigma_i^2)$ then $\bar{X} \sim N(\mu_X, \frac{\sigma_X^2}{n})$

$$\frac{X_1 + \dots + X_n}{n}$$

$$\sum_{i=1}^{n} \frac{1}{n} \mu_X = n\left(\frac{1}{n}\right) \mu_X = \mu_X \qquad \sum_{i=1}^{n} c_i^2 \sigma_X^2 = \sum_{i=1}^{n} \frac{1}{n^2} \sigma_X^2 = n\left(\frac{1}{n^2} \sigma_X^2\right) = \frac{\sigma_X^2}{n}$$

e.g: Suppose $X \sim N(50, 16) \rightarrow \bar{X} \sim N(50, \frac{16}{n})$

$n = 4 \rightarrow \bar{X} \sim N(50, 4)$
$n = 8 \rightarrow \bar{X} \sim N(50, 2)$
$n = 16 \rightarrow \bar{X} \sim N(50, 16)$

$\bar{X} \rightarrow \mu_X$ as $n \rightarrow \infty$

Theorem $\rightarrow$ Let $X_1, X_2 \dots X_n$ be a random sample from $X \sim N(\mu_X, \sigma_X^2)$ then
(a) $\bar{X}$ and S^2 are independent (b) $\frac{(n-1)S^2}{\sigma_X^2} \sim \chi^2(n-1)$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Proof $\rightarrow$ (a) Step 1: Show that $\bar{X}$ and $(X_1 - \bar{X}), \dots (X_{n-1} - \bar{X})$ are independent

Let $Y_n = \bar{X}$, $Y_i = X_i - \bar{X}$, $i = 1, 2 \dots n-1$

↳ show that Y_n and $Y_1, \dots Y_{n-1}$ are independent

$$f_{Y_1 \dots Y_n}(y_1, \dots y_n) = f_{X_1 \dots X_n}(x_1 \dots x_n) \frac{1}{\left|\frac{\partial(y_1, \dots y_n)}{\partial(x_1, \dots x_n)}\right|}$$

inde $\sim N(\mu_X, \sigma_X^2)$

$$= \frac{1}{(\sqrt{2\pi}\sigma_X)^n} e^{\frac{-1}{2\sigma_X^2} \sum_{i=1}^{n} (x_i - \mu_X)^2} \frac{1}{\begin{vmatrix} 1 - 1/n & -1/n & \dots & -1/n \\ -1/n & 1 - 1/n & \dots & -1/n \\ \vdots & & & \vdots \\ -1/n & -1/n & \dots & 1 - 1/n \end{vmatrix}} \dots$$

$$y_1 = x_1 - \bar{x} = x_1 - \frac{x_1 + \dots + x_n}{n}$$

$$= \frac{1}{c} e^{\frac{-1}{2\sigma_X^2}\sum_{i=1}^{n}(x_i-\mu_X)^2}$$

need to use Square Error for sample formula:

$$\boxed{\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2}$$

specific #s
not random variable

Proof →

$$\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}((x_i-\bar{x})+(\bar{x}-\mu))^2$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})^2 + 2\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})(\bar{x}-\mu)}_{0} + \sum_{i=1}^{n}(\bar{x}-\mu)^2$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})^2 + 0 + n(\bar{x}-\mu)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2$$

Q.E.D!

cont.

$$\frac{1}{c} e^{\frac{-1}{2\sigma_X^2}\left(\sum_{i=1}^{n-1}\underbrace{(x_i-\bar{x})^2}_{y_i^2} + (x_n-\bar{x})^2 + \underbrace{n(\bar{x}-\mu)^2}_{ny_n^2}\right)}$$

$$y_1+y_2+\dots+y_{n-1}+(x_n-\bar{x})=0 \rightarrow x_n-\bar{x} = -\sum_{i=1}^{n-1} y_i$$

$$= \frac{1}{c} e^{\frac{-1}{2\sigma_X^2}\left(\sum_{i=1}^{n-1} y_i^2 + \left(\sum_{i=1}^{n-1} y_i\right)^2 + ny_n^2\right)}$$

$$= \frac{1}{c} e^{\frac{-1}{2\sigma_X^2}\left(\sum_{i=1}^{n-1} y_i^2 + \left(-\sum_{i=1}^{n-1} y_i\right)^2\right)} e^{\frac{-1}{2\sigma_X^2} ny_n^2}$$

∴ $y_1, y_2 \dots y_{n-1}$ are independent (since we can write them separately)

Step 2 $\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{i=1}^{n-1} y_i^2 + \left(-\sum_{i=1}^{n-1} y_i\right)^2$ and $Y_n = \bar{X}$ are independent

∴ S^2 and $\bar{X}$ are independent Q.E.D!

Wednesday 4/24

(b) $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 \rightarrow$ Recall: $\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2$

$$\therefore \underbrace{\sum_{i=1}^{n}\left(\frac{X_i-\mu_X}{\sigma_X}\right)^2}_{A\sim\chi^2(n)} = \underbrace{\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma_X}\right)^2}_{B} + \underbrace{\left(\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}}\right)^2}_{C\sim\chi^2(1)}$$

B, C independent

$$M_A(t) = M_B(t).\ M_C(t) \longrightarrow \frac{1}{(1-2t)^{n/2}} = M_B(t).\ \frac{1}{(1-2t)^{1/2}}$$

$$\therefore M_B(t) = \frac{1}{(1-2t)^{n/2-1/2}} = \frac{1}{(1-2t)^{n-1/2}} \rightarrow B\sim\chi^2(n-1)$$

$$\therefore (n-1)\frac{S^2}{\sigma_X^2} \sim \chi^2(n-1)$$ Q.E.D!

Conclusions

Let $X_1, X_2 \dots X_n$ independent identically distributed $X_i \stackrel{d}{=} X \sim N(\mu_X, \sigma_X^2)$

(1) $U = \sum_{i=1}^{n}\left(\frac{X_i-\mu_X}{\sigma_X}\right)^2 \sim \chi^2(n)$

(2) $W = \frac{(n-1)S^2}{\sigma_X^2} = \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma_X}\right)^2 \sim \chi^2(n-1)$

(3) $\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}} \sim N(0,1)$

(4) $\frac{\bar{X}-\mu_X}{S/\sqrt{n}} \sim t(n-1)$ $\rightarrow = \dfrac{\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma_X^2}/n-1}} = \frac{N(0,1)}{\chi^2(n-1)} \sim t(n-1)$

Def

Let $Z\sim N(0,1)$, $U\sim\chi^2(r)$ } Z, U independent then $T = \frac{Z}{\sqrt{U/r}}$ is t-distribution w/ param r $(T\sim t(r))$

$T(r) \rightarrow Z$ as $r\rightarrow\infty$

5.6 Central Limit Theorem

Theorem

Let $X_1, X_2 \dots X_n$ be a random sample of size n from the distribution of X then

$\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}} \rightarrow Z\sim N(0,1)$ in distribution as $n\rightarrow\infty$

$P\left(a < \frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}} < b\right) \rightarrow P(a<Z<b)$ as $n\rightarrow\infty$ for all $-\infty\le a<b\le\infty$

Friday 4/26

5.7 Approximations for Discrete Distributions

Let $X_1, X_2 \dots X_n$ be independent identical distributed $X_i \overset{d}{=} X \sim Ber(p)$, $0<p<1$

Let $Y = X_1 + X_2 + \dots + X_n$ then $Y \sim Bin(n,p)$ then

$$\boxed{\frac{Y-np}{\sqrt{npq}} \to Z \sim N(0,1) \text{ as } n \to \infty}$$

$$\overline{X} = \frac{Y}{n}, \quad \mu_{\overline{X}} = p, \quad \sigma^2_{\overline{X}} = \frac{pq}{n} \longrightarrow CLT: \frac{\overline{X}-p}{\sqrt{\frac{pq}{n}}} \to Z \sim N(0,1)$$

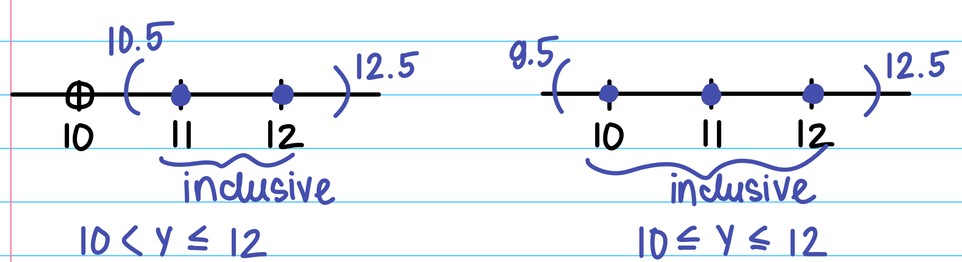
$$\text{OR } \frac{Y-np}{\sqrt{npq}} = \frac{\frac{Y}{n}-p}{\sqrt{\frac{pq}{n}}} \to Z \sim N(0,1) \text{ as } n \to \infty$$

5.7.1 Suppose $Y \sim Bin(25, \frac{1}{2})$. Find approx. value of

correction

$$P(10 < Y \le 12) = P(10.5 < Y < 12.5) = P\left(\frac{10.5-np}{\sqrt{npq}} < \frac{Y-np}{\sqrt{npq}} < \frac{12.5-np}{\sqrt{npq}}\right)$$

$$= P(-0.8 < Z < 0) = \Phi(0) - \Phi(-0.8) = \dots$$



adjust correction based on range

7.1 Confidence Intervals for Means

Let X be a random variable w/ finite mean μ and finite variance σ^2

Suppose we don't know μ & want to estimate μ

Let $X_1, X_2 \dots X_n$ be a random sample of size n from the distribution of X

↳ want to use $\overline{X}$ to estimate μ

I. μ known, σ^2 unknown

(1) $X \sim N(\mu, \sigma^2)$ $100(1-\alpha)\%$ CI is $\left[\overline{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$

(2) Any distribution CI $= \left[\overline{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$ if $n \geq 30$

II μ unknown, σ^2 unknown

(3) $X \sim N(\mu, \sigma^2)$ $\quad CI = \left[\bar{X} - t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right]$ for any n

(4) Any distribution $CI = \left[\bar{X} - Z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \bar{X} + Z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right]$ if $n \geq 30$

Def → Let $A = A(X_1, X_2 \dots X_n)$ and $B = B(X_1, X_2 \dots X_n)$ be statistics from a random sample $X_1, X_2 \dots X_n$ for X

Let $a = A(x_1, x_2 \dots x_n)$ and $b = B(x_1, x_2 \dots x_n)$

$[a, b]$ is called a $100(1-\alpha)\%$ confidence interval for μ if $P(A \leq \mu \leq B) \geq 1-\alpha$

From (1):
$$P\left(\bar{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = P\left(|\bar{X} - \mu| \leq Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$
$$= P\left(\frac{|\bar{X}-\mu|}{\frac{\sigma}{\sqrt{n}}} \leq Z_{\frac{\alpha}{2}}\right) = P\left(|Z| \leq Z_{\frac{\alpha}{2}}\right) \geq 1-\alpha$$ Q.E.D!

From (3):
$$P\left(\bar{X} - t_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right) = P\left(|\bar{X} - \mu| \leq t_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right)$$
$$= P\left(\frac{|\bar{X}-\mu|}{\frac{S}{\sqrt{n}}} \leq t_{\frac{\alpha}{2}}\right) = P\left(|Z| \leq t_{\frac{\alpha}{2}}\right) \geq 1-\alpha$$ Q.E.D! *

From (4): For any dist. X w/ finite μ, σ^2: $\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} \to Z \sim N(0,1)$ if $n \to \infty$

III One-sided CI $= \left[\bar{X} - Z_{\alpha}\frac{\sigma}{\sqrt{n}}, \infty\right)$ if $n \geq 30$