

Definition of Probability

A function P defined on events of a sample space S is called a probability measure if it satisfies the following axioms:

$$\text{Axiom ①: } 0 \leq P(A) \leq 1$$

$$\text{Axiom ②: } P(S) = 1$$

$$\text{Axiom ③: if } A_1, A_2, \dots \text{ is a collection of disjoint events then } P(A_1 \cup A_2 \cup \dots) \\ = P(A_1) + P(A_2) + P(A_3) + \dots = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

* Note disjoint events means no overlap $A_i \cap A_j = \emptyset$
 $\forall i \neq j$

$$\text{Result ①: } P(\emptyset) = 0$$

Result ②: If A_1, A_2, \dots, A_n is a finite collection of disjoint (mutually exclusive)^{*} events then $P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$

Pf Let $A_i = \emptyset$

$$i = n+1, n+2, \dots$$

$$\text{Then } A_i \cap A_j = \emptyset \quad \forall i \neq j$$

$$\text{By Axiom ③ then } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

$$\bigcup_{i=1}^n A_i = \bigcup_{i=1}^n \emptyset$$

$$P\left(\bigcup_{i=1}^n A_i\right) = P\left(\bigcup_{i=1}^n \emptyset\right) = \sum_{i=1}^n P(\emptyset) = \sum_{i=1}^n 0 + \sum_{i=n+1}^{\infty} P(A_i)$$

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


In particular, if A & B are disjoint then $P(A \cup B) = P(A) + P(B)$

Result: Theorem of total probability (3)

Let A_1, A_2, \dots be a partition of the sample space S

i.e. A_1, A_2, \dots are disjoint & $\bigcup_{i=1}^{\infty} A_i = S$

and let B be any event in S

THEN:

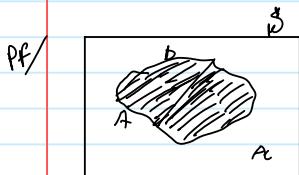


$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots = \sum_{i=1}^{\infty} P(B \cap A_i)$$

* This can be finite as well



$$B = B \cap A_1 \cup B \cap A_2$$



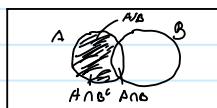
$$B = B \cap A_1 \cup B \cap A_2$$

Note: $B = (B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3) \cup \dots$

& $B \cap A_1, B \cap A_2, \dots$ are disjoint

$$\begin{aligned} P(B) &= P(B \cap A_1) \cup P(B \cap A_2) \cup \dots \\ &= P(B \cap A_1) + P(B \cap A_2) + \dots \quad (\because B \cap A_i's \text{ are disjoint}) \\ &= \sum_{i=1}^{\infty} P(B \cap A_i) \end{aligned}$$

Definition $A \setminus B = A \cap B^c$



Result ④:

For any two events in \mathcal{S}

$$P(A \setminus B) = P(A) - P(A \cap B) = P(A) - P(AB)$$

PF/ Note $A = (A \setminus B) \cup AB$ & $A \setminus B$ and AB are disjoint



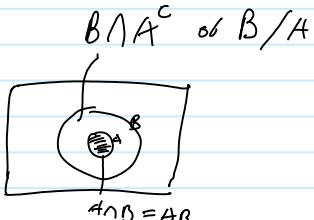
$$\begin{aligned} P(A) &= P((A \setminus B) \cup AB) \\ &= P(A \setminus B) + P(AB) \dots \end{aligned}$$

$$P(A \setminus B) = P(A) - P(AB) \quad (\because A \setminus B \text{ and } AB \text{ are disjoint})$$

Result ⑤:

If $A \subseteq B$, then $P(A) \leq P(B)$

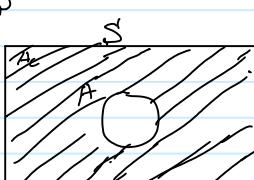
PF/
$$\begin{aligned} P(B \setminus A) &= P(B) - P(AB) \\ &= P(B) - P(A) \geq 0 \\ \Rightarrow P(A) &\leq P(B) \end{aligned}$$



Result 6: For any event in \mathcal{S}

$$P(A^c) = 1 - P(A)$$

PF/
$$\begin{aligned} A^c &= \mathcal{S} - A \\ P(A^c) &= P(\mathcal{S} \setminus A) \\ &= P(\mathcal{S}) - P(SA) \\ &= 1 - P(A) \leftarrow \text{by axiom ②} \end{aligned}$$



Result 7

For any two events $A \& B$ in \mathcal{S}

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

PF) Note that $A \cup B = A \cup (B \setminus A)$ & $A \& B \setminus A$ are disjoint

$$\begin{aligned} P(A \cup B) &= P(A \cup (B \setminus A)) = P(A) + P(B \setminus A) \\ &= P(A) + P(B) - P(AB) \end{aligned}$$



Lecture 01

Probability

Random Experiments: Experiments where the event is random

Let A be an event in a random experiment

Let $r_n(A)$ be the number of times event A occurred in n trials of the experiment, then the probability of A is:

$$P(A) = \lim_{n \rightarrow \infty} \frac{r_n(A)}{n} = \text{relative frequency definition of probability}$$

Events

Events are a subset of a sample space

$$\text{Ex/ } S = \{1, 2\}$$

$$\text{Events} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}, \emptyset \text{ or } \emptyset\}$$

\emptyset = empty set $\emptyset \subset S$ = \emptyset
 \emptyset = not set

A full set is a subset of a set as well as an empty space

$A \cup B$ = elements in A or B (or both)*

$A \cap B$ = elements common to both A & B

$$\text{Ex/ } A = \{1, 2, 3\} \quad B = \{3, 4\}$$

$$A \cap B = \{3\}$$

$$A \cup B = \{1, 2, 3, 4\}$$

* Note don't repeat elements in sets

$$\{1, 2, 3, 3, 4\} \times \{1, 2, 3, 4\} \checkmark$$

Formal definition of Probability

Sample Space (S): The set of all possible outcomes in an experiment

Ex/ ① Tossing coin experiment: $S = \{H, T\}$

② 6 sided die experiment: $S = \{1, 2, 3, 4, 5, 6\}$

③ Tossing 2 coins experiment: $S = \{HH, HT, TH, TT\}$

* Assume different coins when it is not specified

Results

$$\textcircled{1} \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$\textcircled{2} \quad (A \cap B)^c = A^c \cup B^c$$

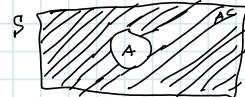
$$\textcircled{3} \quad (A \cup B)^c = A^c \cap B^c$$

$$\textcircled{4} \quad A \cup \emptyset = A$$

$$\textcircled{5} \quad A \cap \emptyset = \emptyset$$

A^c = elements in sample space but not in A

A^c = complement



$$A \cap B = AB$$

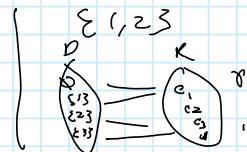
Definition of Probability

Let S be a sample space A is a function P defined on events is a probability if

Axioms $\textcircled{1} \quad 0 \leq P(A) \leq 1$

$\textcircled{2} \quad P(S) = 1$

$\textcircled{3} \quad$ if A_1, A_2, \dots is countable collection of disjoint events then $P(A_1 \cup A_2 \cup A_3 \dots) = P(A_1) + P(A_2) + P(A_3) \dots$
 $= P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$



Last hour

Counting given the digits $\overbrace{[1][2]\dots[9]}^n$

how many 4 digit numbers can be formed?

$$\text{Answer: } \frac{9 \cdot 8 \cdot 7 \cdot 6}{1 \cdot 2 \cdot 3 \cdot 4} = 9 \cdot 8 \cdot 7 \cdot 6 \cdot \frac{(5 \cdot 4 \cdot 3 \cdot 2 \cdot 1)}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

$$= \frac{9!}{5!} = \frac{9!}{(9-5)!} = \frac{9!}{4!}$$

Combinations: The number of combination of k objects from n objects = $\frac{n!}{k!(n-k)!}$

Note: The number of ways of dividing n objects into two groups of sizes n_1 & n_2

$$= \frac{n!}{n_1! n_2!}$$

Permutation with repetition: $\frac{n^r}{r!}$

Permutation without repetition: $\frac{n!}{(n-r)!}$

Combination with repetition: $\frac{(n+r-1)!}{r!(n-1)!}$

Combination without repetition: $\frac{n!}{(n-r)! r!}$

The number of ways of dividing n objects into groups of sizes

$$n_1, n_2, \dots, n_r \text{ is: } \frac{n!}{n_1! n_2! \dots n_r!} = \binom{n}{n_1, n_2, \dots, n_r}$$

Ex/ A student prepares for an exam by studying a list of 10 questions. She can answer only 6 questions. For the exam the instructor selects 5 questions from the 10 questions at random. What is the probability that she can solve all the 5 questions on the exam

$$\begin{array}{ccccccc} 10 & 9 & 8 & 7 & 6 & & \\ \hline & & & & & & \end{array} \quad \frac{\binom{6}{5} \times \binom{4}{0}}{\binom{10}{5}}$$

$$\frac{n!}{k!(n-k)!} = \frac{10!}{(10-5)!} = \frac{10!}{5!} \rightarrow \# \text{ of possibilities in the exam}$$

Ex A deck of 52 cards contains 12 picture cards. If 52 cards are distributed at random among 4 people, such that each player gets 13 cards. What is the probability that each player gets 3 picture cards.

$$\binom{52}{13 \ 13 \ 13 \ 13} = \frac{52!}{13! \ 13! \ 13! \ 13!}$$

$$\binom{12}{3 \ 3 \ 3 \ 3} \times \binom{40}{10 \ 10 \ 10 \ 10} = \frac{12!}{3! \ 3! \ 3! \ 3!} \times \frac{40!}{10! \ 10! \ 10! \ 10!}$$

$$\frac{12!}{3! \ 3! \ 3! \ 3!} \times \frac{40!}{10! \ 10! \ 10! \ 10!}$$

$$\frac{12!}{3!7!2!1!} \times \frac{40!}{10!16!10!10!}$$

$$= \underline{\underline{52!}}$$

$$\underline{\underline{13!13!13!13!}}$$

Ex/ One of my classes has 60 students. I want to divide this class into two groups, each with 30 students. John and Jane are two friends in this class. what is the probability that John & Jane will be in the same class?

$$\frac{60!}{30!30!} = \frac{(58)!}{28!28!} \cdot \frac{(58)!}{30!}$$

Conditional Probability (C.S.)

Def" Let A & B be any two events in sample space S Then the conditional probability of A given B is: $P(A|B) = \frac{P(A \cap B)}{P(B)}$, if $P(B) > 0$

Ex/ For two events A & B, $P(A)=0.5$, $P(B)=0.3$, $P(AB)=0.1$

$$\text{Find: a) } P(A|B) \Rightarrow P(A|B) = \frac{P(AB)}{P(B)} = \frac{0.1}{0.3}$$

$$\text{b) } P(B|A) \Rightarrow \frac{P(AB)}{P(A)} = \frac{0.1}{0.5}$$

Note $P(A|B) \neq P(B|A)$

$$P(AB) = P(BA) = P(A \cap B) = P(B \cap A)$$

$$\text{c) } P(A|A \cup B) = \frac{P(A \cap A \cup B)}{P(A \cup B)} = \frac{P(A)}{P(A) + P(B) - P(A \cap B)} = \frac{0.5}{0.5 + 0.3 - 0.1} = \frac{0.5}{0.7} = \frac{5}{7}$$

$$\text{d) } P(A|AB) = \frac{P(A \cap AB)}{P(AB)} = \frac{P(AB)}{P(AB)} = 1$$

$$\text{e) } P(A \cap B | A \cup B) = \frac{P(A \cap B \cap A \cup B)}{P(A \cup B)} = \frac{P(A \cap B)}{P(A \cup B)} = \frac{P(A \cap B)}{P(A) + P(B) - P(AB)}$$

Last hour:

Conditional Probability:

Definition: If A & B are two events in \mathcal{S} s.t $P(B) > 0$

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

Note: conditional probability is a probability measure.

$$\begin{aligned} P(A \cap B) &\leq B \\ 0 \leq P(A \cap B) &\leq P(B) = 1 \end{aligned}$$

$$0 \leq P(A|B) \leq 1 \quad \underline{\text{Axiom 1}}$$

Axiom 2

$$P(S|B) = \frac{P(S \cap B)}{P(B)} \quad (\text{definition})$$

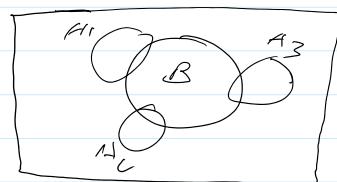
$$P(B|B)$$

$$= 1$$

Axiom 3

Let A_1, A_2, \dots be a countable collection of disjoint events.

$$P(A_1 \cup A_2 \cup \dots | B) = \frac{P((\overbrace{A_1 \cup A_2 \cup \dots}^{\text{disjoint}}) \cap B)}{P(B)} \dots (\text{Def})$$



$$\begin{aligned} &= \frac{P(A_1 B \cup A_2 B \cup \dots \cup A_n B)}{P(B)} \\ &= \frac{P(A_1 B) + P(A_2 B) + \dots}{P(B)} \end{aligned}$$

($A_1 B, A_2 B, \dots$ are disjoint)

$$= \frac{P(A_1 B)}{P(B)} + \frac{P(A_2 B)}{P(B)} + \dots$$

$$= P(A_1|B) + P(A_2|B) + \dots$$

ie $P(\cdot|B)$ is a probability measure

The Theorem of Total Probability

Conditional Probability version.

Let A_1, A_2, \dots be a partition of S

partition means:

- (1) A_1, A_2, \dots are disjoint
- (2) $\bigcup_{i=1}^{\infty} A_i = S$ (can be finite)

and B be any event in S . Then

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \dots \\ &= \sum_{i=1}^{\infty} P(B|A_i)P(A_i) \end{aligned}$$

This is because $P(A_1|B) = P(B|A_1)P(A_1)$

Theorem Bayes Theorem

Let A & B be two events in S ($P(A) > 0, P(B) > 0$)

$$P(B|A_1) = \frac{P(A_1|B)}{P(A_1)}$$

$$P(A, B) = P(B|A_1)P(A_1)$$

$$\text{Then } P(A|B) = \frac{P(A)}{P(B)} \times P(B|A)$$

PF

$$P(A|B) \stackrel{\text{def}}{=} \frac{P(AB)}{P(B)}$$

$$\left| \begin{array}{l} P(B|A) = \frac{P(AB)}{P(A)} \\ P(AB) = P(B|A)P(A) \end{array} \right.$$

$$= \frac{P(A)}{P(B)} P(B|A) \quad \boxed{\text{}}$$

A population of voters consists of 40% republicans and 60% democrats. 30% of republicans and 50% of democrats favor a particular election issue. A voter selected at random

In population of 1000 randomly chosen voters, 40% are republicans and 60% democrats. 30% of republicans and 50% of democrats favor a particular election issue. A voter selected at random is found to favor the issue. Find the conditional probability that he is a democrat.

$$P(R) = 0.4$$

$$P(D) = 0.6$$

$$P(S) = P(R \cup D) = 1$$

$$P(F|R) = 0.3$$

$$P(F|D) = 0.5$$

$$\textcircled{1} \quad P(D|F) = \frac{P(D)}{P(F)} \times P(F|D)$$

$$\begin{aligned} \textcircled{2} \quad P(F) &= P(F|R)P(R) + P(F|D)P(D) \\ &= 0.3(0.4) + 0.5(0.6) \\ &= 0.12 + 0.3 \\ &= 0.42 \end{aligned}$$

$$\textcircled{3} \quad P(D|F) = \frac{0.6}{0.42} \times 0.5$$

1.5.2 Independent events

Let A & B be two events in \mathcal{S} , we say A & B are independent if $P(A \cap B) = P(A) \times P(B)$

Note: if $P(B) > 0$

if $P(B) > 0$

$A \perp\!\!\!\perp B$ if $P(A|B) = P(A)$
↓
independent if

$$\left| \begin{array}{l} P(A|B) = \frac{P(AB)}{P(B)} \\ \text{if independent} \\ P(AB) = P(A)P(B) \\ P(A|B) = P(A) \end{array} \right.$$

Example: Roll a 6 sided die ... $S = \{1, 2, 3, 4, 5, 6\}$

$$A = \text{observe an odd number} = \{1, 3, 5\}$$

$$B = \text{observe an even number} = \{2, 4, 6\}$$

$$C = \text{observe a number} < 3 = \{1, 2\}$$

Q. a) Are A & B independent

$$\begin{array}{lll}
 P(A) = \frac{1}{2} & AB = \emptyset & P(AB) = P(A) \times P(B) \\
 P(B) = \frac{1}{2} & P(AB) = 0 & 0 = 0.5 \times 0.5 \\
 P(C) = \frac{1}{3} & & 0 \neq 0.25
 \end{array}$$

$\therefore A \& B$ are not independent.

b) Are $A \& C$ independent?

$$AC = \{1\} \quad P(AC) = \frac{1}{6} = P(A) \times P(C) = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$$

$\therefore A \perp C$

Note: If A and B are disjoint events with $P(A) > 0$ and $P(B) > 0$, Then $A \cap B$ can't be independent

Reason: if $A \& B$ are disjoint. $AB = \emptyset$ and so the

$$P(AB) = 0 \neq P(A)P(B)$$

Result $A \perp B \Rightarrow A \perp B^c$

$$\begin{aligned}
 P: P(A \cap B^c) &= P(A \cap B) \\
 &= P(A) - P(AB) \\
 &= P(A) - P(A)P(B) \\
 &= P(A)[1 - P(B)] \\
 &= P(A)P(B^c)
 \end{aligned}$$

$$\left. \begin{aligned}
 P(AB) &= P(A)P(B) \\
 \text{To prove:} \\
 P(A \cap B^c) &= P(A)P(B^c)
 \end{aligned} \right\}$$

$\therefore A \perp B^c \quad \square$

Ex/ show that $A \perp B \Rightarrow A^c \perp B^c$

Definition: Events A_1, A_2, A_3, \dots are said to be independent if $P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \times P(A_{i_2}) \times \dots \times P(A_{i_k})$ for each subcollection of events $A_{i_1}, A_{i_2}, \dots, A_{i_k}$

$A_{i_1}, A_{i_2}, A_{i_3}, A_{i_4}, A_{i_5}$

$$P(A_2 \cap A_5) = P(A_2)P(A_5)$$



Ex/ Roll a 4 sided (tetrahedral) die.
 $S = \{1, 2, 3, 4\}$

$$\text{Let } A = \{1, 2\} \quad B = \{1, 3\} \quad C = \{1, 4\}$$

(a) Are A & B independent?

$$P(AB) = P(A) \times P(B)$$

$$\frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$$

$$\frac{1}{4} = \frac{1}{4}$$

$$\therefore A \perp B$$

$$\left. \begin{array}{l} P(A) = P(B) = P(C) = \frac{2}{4} = \frac{1}{2} \\ P(AB) = \frac{1}{4} \\ P(BC) = \frac{1}{4} \end{array} \right\}$$

(b) Are $B \perp C$?

(c) Are $A \perp C$

$$P(BC) = P(B) \times P(C)$$

$$\frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$$

$$\frac{1}{4} = \frac{1}{4}$$

$$P(AC) = P(A) \times P(C)$$

$$\frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$$

$$\frac{1}{4} = \frac{1}{4}$$

A, B, C are pairwise independent

(d) Are A, B, C independent

$$ABC = \{1\}$$

$$P(ABC) = \frac{1}{4}$$

$$P(A) \times P(B) \times P(C) = P(ABC)$$

$$\frac{1}{8} = P(ABC)$$

$\frac{1}{8} \neq \frac{1}{4} \therefore A, B$ and C are not independent.

Note: Pairwise independence does not imply independence.

Continuity of Probability

Definition:

Let A_1, A_2, \dots be a collection of sets

We say A_n increases A ($A_n \uparrow A$) if $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$
and $A = \bigcup_{n=1}^{\infty} A_n$

We write $\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n$

Definition: we can say $A_n \uparrow A$ if $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$

f is c/s at a if
 $\lim_{x \rightarrow a} f(x) = f(a)$

Recall

f is continuous at a
if for every sequence $\{x_n\}$
 $\rightarrow a$ such that $\lim_{n \rightarrow \infty} x_n = a$

$$f(a) = \lim_{n \rightarrow \infty} f(x_n)$$

Definition: we say $\{A_n\} \downarrow A$ if $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$

and

$$A = \bigcap_{n=1}^{\infty} A_n$$

$$\text{We write } \lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$$

$$\left\{ \begin{array}{l} f(\alpha) = \lim_{n \rightarrow \infty} f(x_n) \\ \int \left(\lim_{n \rightarrow \infty} x_n \right) = \lim_{n \rightarrow \infty} (x_n) \end{array} \right.$$

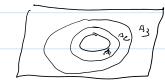
Theorem: 1.6.1 Continuity of Probability

If $\{A_n\} \uparrow A$ or $\downarrow A$, then

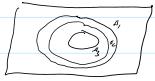
$$P(A) = \lim_{n \rightarrow \infty} P(A_n)$$

Last hour

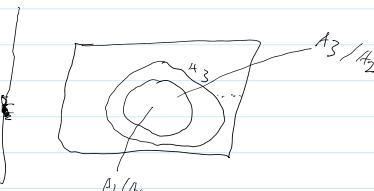
- $\sum A_n \uparrow A$ if $A_1 \subseteq A_2 \subseteq \dots$
and $A = \bigcup_{n=1}^{\infty} A_n$ (we call this $\lim_{n \rightarrow \infty} A_n$)



- $\sum A_n \downarrow A$ if $A_1 \supseteq A_2 \supseteq \dots$
and $A = \bigcap_{n=1}^{\infty} A_n$

(we call this $\lim_{n \rightarrow \infty} A_n$)Theorem 1.6.1 Continuity theorem of probabilityif $\sum A_n \uparrow A$ or $\sum A_n \downarrow A$ then $P(A) = \lim_{n \rightarrow \infty} P(A_n)$ Proof case 1: $\sum A_n \uparrow A$

$$A = \bigcup_{n=1}^{\infty} A_n = A_1 \cup (\bigcup_{i=2}^{\infty} (A_i \setminus A_{i-1}))$$



$$P(A) = P(A_1 \cup (\bigcup_{i=2}^{\infty} (A_i \setminus A_{i-1})))$$

$$= P(A_1) + P(\bigcup_{i=2}^{\infty} (A_i \setminus A_{i-1}))$$

 $\because A_i = \bigcup_{j=2}^{\infty} (A_j \setminus A_{i-1})$ are disjoint

$$= P(A_1) + \sum_{i=2}^{\infty} P(A_i \setminus A_{i-1}) \dots$$

 $(A_i \setminus A_{i-1}$ are disjoint)

$$= P(A_1) + \lim_{n \rightarrow \infty} \sum_{i=2}^{\infty} P(A_i \setminus A_{i-1}) \quad * \text{Remember: } P(A \setminus B) = P(A) - P(AB)$$

$$= P(A_1) + \lim_{n \rightarrow \infty} \sum_{i=2}^{\infty} [P(A_i) - P(A_i \setminus A_{i-1})] \quad * \ P(A_i \setminus A_{i-1}) = P(A_i) \text{ because } A_i \subseteq A_{i-1}$$

$$= P(A_1) + \lim_{n \rightarrow \infty} \sum_{i=2}^n [P(A_i) - P(A_{i-1})]$$

$$\left| \sum_{i=1}^n [P(A_i) - P(A_{i-1})] \right| = P(A_2) - P(A_1) + P(A_3) - P(A_2) + P(A_4) - P(A_3) \dots + P(A_n) - P(A_{n-1})$$

$$= P(A_1) + \lim_{n \rightarrow \infty} P(A_n) - P(A_1)$$

$$\lim (a_n + b_n) = \lim(a_n) + \lim(b_n)$$

$$= \lim_{n \rightarrow \infty} P(A_n)$$

Case 2 If $\sum A_n \downarrow A = \bigcap_{n=1}^{\infty} A_n$ Let $B_n = A_n^c$ then A^c

$$B_n \uparrow \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n^c = (\bigcap_{n=1}^{\infty} A_n)^c \quad \text{De Morgan's law}$$

i.e. $\sum B_n \uparrow A^c$

$$\therefore \text{By case 1: } P(A^c) = \lim_{n \rightarrow \infty} P(B_n)$$

Remember: $P(A^c) = 1 - P(A)$

$$1 - P(A) = \lim_{n \rightarrow \infty} P(A_n^c)$$

$$1 - P(A) = \lim_{n \rightarrow \infty} (1 - P(A_n))$$

$$\lim_{n \rightarrow \infty} P(A_n) = P(A)$$

$$P(A) = \lim_{n \rightarrow \infty} P(A_n)$$

Example: 1.6.4 p30 / Suppose $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \frac{2+e^{-i}}{6}$, $i = 1, 2, \dots$

Find $P(\Sigma A_i)$?

Solution: $A_n \downarrow \Sigma A_i$!



$$\text{Ex/ } S = \{1, 2, 3, \dots\} \quad P(\Sigma S) = \left(\frac{1}{2}\right)^{\infty}$$

Find $P(\Sigma S, 6, \dots)$ using continuity of probability and finite additivity

Solution Define $A_n = \{S, 6, \dots, n\}$ then $\bigcap A_n = \emptyset$
using continuity of probability - $P(\Sigma S, 6, \dots) = \lim_{n \rightarrow \infty} P(A_n)$

$$A_n = \{S, 6, \dots, n\} \quad \text{← } n-4 \text{ elements in } A$$

$$\begin{aligned} P(A_n) &= P(S) + P(S, 6) + \dots + P(S, 6, \dots, n) \\ &= \underbrace{\left(\frac{1}{2}\right)^5}_{a} + \underbrace{\left(\frac{1}{2}\right)^6}_{r} + \dots + \underbrace{\left(\frac{1}{2}\right)^n}_{r^n} \\ &= \left(\frac{1}{2}\right)^5 \underbrace{\left[1 - \left(\frac{1}{2}\right)^{n-4}\right]}_{1-r} \end{aligned}$$

$$\left| \begin{array}{l} \text{Geometric sequence/series} \\ S_n = \frac{a(1-r^n)}{1-r} \end{array} \right.$$

$$= \left(\frac{1}{2}\right)^5 \left[1 - \left(\frac{1}{2}\right)^{n-4}\right]$$

$$\lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^5 \left[1 - \left(\frac{1}{2}\right)^{n-4}\right] = \left(\frac{1}{2}\right)^5 = \frac{1}{16}$$

Last hour (continuity of probability)

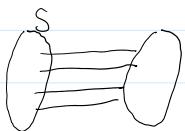
[Midterm * tentative: July 6th]

if $\sum A_n \uparrow A$ or $\sum A_n \downarrow A$ Then $P(A) = \lim_{n \rightarrow \infty} P(A_n)$



Today: Random variables and their distributions (chapter 2.)

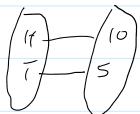
Definition: A random variable (r, v) is a function from the sample space to the set of real numbers (\mathbb{R}).



Example Toss a coin $S = \{H, T\}$

$$x(H) = \$10$$

$$x(T) = \$-5$$



Example 2 Roll a 6 sided die $S = \{1, 2, 3, 4, 5, 6\}$

$$x(S) = \$$$



Example 3 Toss two coins $S = \{HH, HT, TH, TT\}$

$$x(S) = \# \text{ of heads in } S$$

$$x(HH) = 2$$

$$x(HT) = 1$$

$$x(TH) = 1$$

$$x(TT) = 0$$

$$S = \{1, 2, 3, 4, 5, 6\}$$

$$x(S) = \begin{cases} 1 & \text{if } S \in \{HH, HT, TH\} \\ 0 & \text{if } S \in \{TT\} \end{cases}$$

Indicator Function

Indicator Function

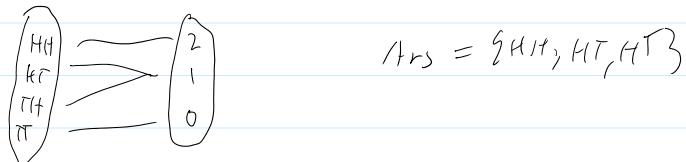
$$I_A(s) \begin{cases} 1 & s \in A \\ 0 & s \notin A \end{cases}$$

Definition

$x = y$ means $x(s) = y(s) \forall s \in S$

Definition For $B \subseteq \mathbb{R}$, $x \in B \Leftrightarrow \{s \mid s \in S \text{ and } x(s) \in B\}$

Find $x \in \{1, 2, 3\}$



$$x_{rs} = \{HHT, HTT, HTH\}$$

Note: $x \in B \subseteq S$ for any $B \subseteq \mathbb{R}$ and so $P(x \in B)$ is defined.

Distribution of a random variable

Definition: The collection of probabilities $P(x \in B)$ for all $B \subseteq \mathbb{R}$ is called the distribution.

Ex/ $S = \{\text{clear, snow, rainy}\}$

$$P(\text{clear}) = 0.5, P(\text{snow}) = 0.2, P(\text{rain}) = 0.3$$

$$x(\text{cl}) = 200$$

$$x(\text{sn}) = 100$$

$$x(\text{ra}) = -50$$

Find the distribution of x ?

Let $B \subseteq \mathbb{R}$

$$\left| \begin{array}{l} B = (140, 210) \\ P(x \in (140, 210)) = P(x=200) \end{array} \right.$$

Let $B \subseteq \mathbb{R}$

$$P(x \in B) = \alpha_1 I_B^{(200)} + \alpha_2 I_B^{(100)} + \alpha_3 I_B^{(-50)}$$

$$\left. \begin{aligned} P &= (140, 20) \\ P(x \in (140, 20)) &= P(x=20) \\ &= f(1) \\ &= 0.5 \end{aligned} \right\}$$

$$B = \{10, 20\}$$

$$\left. \begin{aligned} P(x \in (10, 20)) &= P(x=10 \text{ or } x=20) \\ &= 0.2 + 0.5 \\ &= 0.7 \end{aligned} \right\}$$

$$P(x \in (-60, 500)) =$$

Ex/ Toss a coin twice $S = \{HH, HT, TH, TT\}$

you win

\$4 if you get 2 H's

\$1 if you get 1 H

-\$5 if you get no heads

Let $x = \text{amount of money you win}$. Find
the distribution of x .

$$P(x \in R) = \frac{1}{4} I_R^{(4)} + \frac{1}{2} I_R^{(1)} + \frac{1}{4} I_R^{-5}$$

Discrete Random Variables

Definition: A random variable is discrete if there is a finite or a countably infinite sequence of real numbers x_1, x_2, \dots and a corresponding sequence of positive real numbers p_1, p_2, \dots such that

$$P(x=x_i) = p_i \quad \text{and} \quad \sum p_i = 1$$

Definition: Probability mass function of discrete random variables.

The prob mass function (pmf) of a discrete random variable

X is the function $P_X: \mathbb{R} \rightarrow [0,1]$ defined by
 $P_X(x) = P(X=x), \forall x \in \mathbb{R}$

Ex

x	1	2	3	4	5	6
P_X	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

or $P_X(x) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6$

x	1	2	3	4	5	6
P_X	0.1	0.2	0.2	0.3	0.1	0.1

$$\begin{aligned} P(X \in (4.5, 6.5)) &= \frac{1}{6} + \frac{1}{6} = \\ &= P_X(5) + P_X(6) \\ &= \sum P_X(x) \quad x \in (4.5, 6.5) \end{aligned}$$

$$P(X \in A) = \sum_{x \in A} P(X=x)$$

Important discrete distribution

① Degenerate distribution at c

A random variable X is said to be degenerate at c if

$$P_X(x) \begin{cases} 1 & x=c \\ 0 & x \neq c \end{cases}$$

x	1	2	3	4	5	6
P_X	$\frac{1}{2}$			$\frac{1}{2}$		

② Bernoulli distribution

A random variable X is said to have a Bernoulli distribution with parameter θ (we write $X \sim \text{Bernoulli}(\theta)$) if:

$$P_X(x) \begin{cases} \theta & x=1 \\ 1-\theta & x=0 \end{cases}$$

$$P_x^{(x)} \left\{ \begin{array}{l} \theta \quad x=1 \\ 1-\theta \quad x=0 \end{array} \right.$$

x	0	1
$P_x^{(x)}$	$1-\theta$	θ

①

③ Binomial Distribution

A random variable X is said to have a binomial distribution with parameters $n \in \Theta$ ($X \sim \text{Binomial}(n, \theta)$) if $P_x(x) = \binom{n}{x} p^x (1-\theta)^{n-x}$, $x=0, 1, 2 \dots n$

$$x = \# \text{ of "1's"}$$

$$P_x^{(x)} = P(X=x)$$

$$\sum_{x=0}^n P_x^{(x)} = ?$$

$$\sum_{x=0}^n (P_x^{(x)}) = \sum_{x=0}^n \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

n	0, ..., n
P_x	

$$(a+b)^n = \sum_{x=0}^n a^x b^{n-x}$$

$$\begin{aligned} &= (\theta + 1 - \theta)^n \\ &= 1^n \\ &= 1 \end{aligned}$$

④ Geometric Distribution

A random variable X is said to have a geometric distribution with parameter

$$\theta = (X \sim \text{geometric}(\theta))$$

$$\text{if } P_x^{(x)} = (1-\theta)^x \theta, \quad x=0, 1, 2$$

$X = \# \text{ of failures before the first success}$

$$P_x^{(x)} = P(X=x) = (1-\theta)^x \theta$$

x	0	1
P_x				

$$Q^x : \sum_{n=0}^{\infty} P_x^{(x)} = ?$$

$$\sum_{n=0}^{\infty} (1-\theta)^x \theta = \frac{\theta}{1-(1-\theta)^x}$$

$$\sum_{x=0}^{\infty} p_x^{(\theta)} = \sum_{x=0}^{\infty} (1-\theta)^x \theta = \theta \sum_{x=0}^{\infty} (1-\theta)^x$$

$$= \theta [1 + (1-\theta) + (1-\theta)^2 + (1-\theta)^3 + \dots] \Rightarrow \text{Geometric series}$$

$a = 1$
 $r = (1-\theta)$

$$S_{\infty} = \frac{a}{1-r} = \frac{1}{1-(1-\theta)} = \frac{1}{\theta} = 1$$

$\therefore x \sim P(\theta)$ \uparrow pmf.

(5) Negative Binomial distribution:

A random variable x is said to have a negative binomial distribution with parameters r & θ

$$P_x^{(r)} = \binom{x+r-1}{x} (1-\theta)^x \theta^r$$

$$x = 0, 1, 2, \dots$$

$$Q \quad \text{Is } \sum_{x=0}^{\infty} p_x(x) = 1 ?$$

$$\sum \binom{n+r-1}{r} (1-\theta)^x \theta^r = 1 ?$$

Midterm Test: July 6 (9:00 - 11:00) \rightarrow (7:00 - 9:00 pm)

Rooms AA113 AC223

Last hour: Important discrete functions

① Degenerate distribution $P_x(x) = \begin{cases} 1 & x=c \\ 0 & x \neq c \end{cases}$

x	c
P_x	1

② Bernoulli (θ): $P_x(x) = \begin{cases} \theta & x=0 \\ 1-\theta & x=1 \end{cases}$

x	0	1
P_x	$1-\theta$	θ

③ Binomial (n, θ): $P_x(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad x=0, 1, \dots, n$

x	0, ..., n
P_x	...

④ Geometric (θ): $P_x(x) = (1-\theta)^x \theta$

of failures before the first

success \rightsquigarrow 1 success

⑤ Negative binomial distribution $\rightsquigarrow 1 \leq$ success

$x = \#$ of failures before the r^{th} success

x is negative binomial (r, θ) if $P_x(x) = \binom{x+r-1}{x} (1-\theta)^x \theta^r \quad x=0, 1, 2, \dots$

Q1 $\sum \binom{x+r-1}{x} (1-\theta)^x \theta^r = ?$

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

$$\binom{5}{3} = \frac{5 \times 4 \times 3}{1 \times 2 \times 3}$$

$$(1+t)^n = \sum_{k=0}^n \binom{n}{k} t^{n-k}$$

$$\binom{7}{x} = \frac{n(n-1) \dots (n-(x-1))}{1 \dots 2 \dots x}$$

$$\frac{1}{1+t} = (1+t)^{-1} = 1-t+t^2-t^3 \dots$$

$$(1+t)^{-r} = \sum_{k=0}^{\infty} \binom{-r}{k} t^k$$

$$\binom{x+r-1}{x} = \frac{(x+r-1)(x+r-2) \dots (x+r-1-(x-1))}{1 \dots 2 \dots x}$$

$$= \frac{r(r+1)(r+2) \dots (r+x-1)}{1 \dots 2 \dots x}$$

$$= \underbrace{(-1)^x (-r)(-r-1)(-r-2) \dots (-r-(r-1))}_{1-2 \dots x}$$

$$= (-1)^x \binom{-r}{x}$$

$$\sum_{x=0}^{\infty} \binom{x+r-1}{x} (1-\theta)^x \theta^r = \sum_{x=0}^{\infty} (-1)^x \binom{-r}{x} (1-\theta)^x \theta^r$$

constant

$$= \theta^r \sum_{x=0}^{\infty} (-1)^x \binom{-r}{x} (1-\theta)^x \quad | \quad a^x b^x = (ab)^x$$

$$= \theta^r \sum_{x=0}^{\infty} \binom{-r}{x} (\theta-1)^x \quad | \quad (1+\epsilon)^{-r} = \sum_{x=0}^{\infty} \binom{-r}{x} \epsilon^x$$

$$= \theta^r [1 + \theta - 1]^{-r}$$

$$= \theta^r \theta^{-r} = 1$$

f_x is a pmf

Ex $P(\text{succ}) = 0.6$: Find the probability of at least 2 failures before the third success.

Let x be the # of failures before the third success; $x \sim \text{Neg binomial}$
 $r = 3$
 $\theta = 0.6$

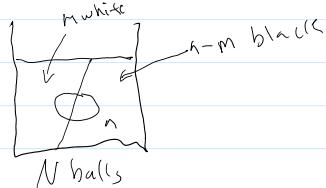
$$P(x \geq 2) = 1 - P(x=1) - P(x=0)$$

$$= 1 - \left[P(x=0) + P(x=1) \right]$$

$$P(x=0) = \binom{3+3-1}{3} (1-0.6)^3 (0.6)^0$$

$$P(x=1) = \binom{3+3-1}{1} (1-0.6)^1 (0.6)^2$$

(6) Hypergeometric distribution



$$f_x(x) = P(x \geq k) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

A random variable x is said to have a hypergeometric distribution if (n, M, N, r) .

A random variable x is said to have a hypergeometric distribution if $x \sim \text{hypergeometric}(N, M, n)$

$$\text{if pmf, } p_x(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

$$\begin{cases} 0 \leq M \\ 0 \leq n \\ n \leq N \\ x \geq 0 \\ n-x \leq N-M \end{cases} \quad x \leq \min(n, M)$$

$$\max(0, n - (N-m)) \leq x \leq \min(n, m)$$

⑦ Poisson distribution

A rv x is said to have a Poisson distribution with parameter λ if

$$p_x(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad x = 0, 1, 2, \dots$$

recall

$$e^\lambda = 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots$$

Term Test: July 6th: 1:00 - 2:00

Rooms: AA12, AC223

Last hour: Important discrete distributions

① Degenerate Distribution $P_{X^{(0)}} = \begin{cases} 1 & x=c \\ 0 & x \neq c \end{cases}$ ② Bernoulli (θ) distribution $P_x(x) = \theta^x (1-\theta)^{1-x}$, $x=0, 1$ ③ Binomial (n, θ) $P_x(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$, $x=0, 1, \dots, n$ ④ Geometric θ distribution [number of failures before the 1st success]

$$P_x(x) = (1-\theta)^x \theta, \quad x=0, 1, 2, \dots$$

⑤ Negative binomial distribution [number of failure before rth success]

$$P_x(x) = \binom{x+r-1}{x} (1-\theta)^x \theta^r$$

⑥ Hypergeometric distribution (hypergeometric (N, M, n))

$$P_x(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, \quad \max(0, n-(m-n)) \leq x \leq \min(n, m)$$

⑦ Poisson distribution ($\lambda \sim \text{Poisson}(\lambda)$) ← questions for
Poisson with Gr
told but
not for the othersPoisson process and queues

- Consider a situation involving a server [cashier, Doctor]

- We assume that the units arrive in a random fashion.

- In situations like this, if X is the number of arrivals in an interval of length t , X typically has a Poisson (λt), where λ is the rate of arrivals [i.e. the average number of arrivals per unit time].Ex) Telephone calls arrive at a help line at a rate 2 per minute [$\lambda=2$]a) what is the probability that 5 calls arrive in the next 2 minutes [$t=2$])

$$X \sim P_0 (\lambda t = 2 \times 2 = 4), \quad P(x=5) = \frac{e^{-4} \lambda^5}{5!}$$

b) what is the probability no calls during the next 10^{min}

$$\lambda = \# \text{ of calls in the next } 10 \text{ min} \sim P_0(\lambda | t) = f_{\lambda}(20)$$

$$P(x=0) = e^{-20} \frac{(20)^0}{0!} = e^{-20}$$

c) What is the probability that 5 calls in the next 2 minutes and 5 more calls in the following 2 minutes?

$$(0.16) \times (0.16) = 0.16^2$$

Continuous distribution

A random variable X is defined to be continuous if $P(X=x)=0 \forall x \in \mathbb{R}$

$$\text{if } P(v=x)=0 \quad \forall v \in \mathbb{R}$$

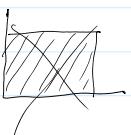
Definition

A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is called a density function if it satisfies

① $f(x) \geq 0 \quad \forall x \in \mathbb{R}$



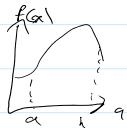
② $\int_{-\infty}^{\infty} f(x) dx = 1$



Definition: Absolutely continuous random variables

A random variable X is said to be absolutely continuous if \exists a density function f s.t.

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$



$$\text{Note: } P(x=x) = P(x \leq X \leq x) = \int_x^x f(x) dx = 0$$

Note: A random variable with absolute continuous \Rightarrow continuous

Note converse is false ($\text{cts} \not\Rightarrow \text{abs}$, $\text{cts} \Rightarrow \text{abs}$)

Ex/ a function that is continuous but not abs
is uniform distribution on center set.

Ex/ An absolute continuous (abs cts) r.v. x has probability density function (pdf).

$$f(x) = kx, \quad 1 \leq x \leq 5 \\ = 0$$

(a) Find the value of k

(b) Find $P(2 \leq x \leq 3)$

$$(a) \int_{-\infty}^{\infty} f(x) dx = 1 \quad \text{is} \quad \int_{-\infty}^1 f(x) dx + \int_1^5 f(x) dx + \int_5^{\infty} f(x) dx = 1$$

$\underbrace{\hspace{1cm}}$ $\underbrace{\hspace{1cm}}$ $\underbrace{\hspace{1cm}}$

$$\int_1^5 f(x) dx = \left[\frac{kx^2}{2} \right]_1^5 = 1$$

$$k \left[\frac{25}{2} - \frac{1}{2} \right] = 1$$

$$k[12] = 1$$

$$k = \frac{1}{12}$$

$$(b) P(2 \leq x \leq 3) = \int_2^3 f(x) dx \\ = \int_2^3 \frac{1}{12} x dx = \frac{1}{12} \left[\frac{x^2}{2} \right]_2^3 \\ = \frac{1}{12} \left[\frac{9}{2} - \frac{4}{2} \right] = \frac{1}{12} \times \frac{5}{2} = \frac{5}{24}$$

Ex 2 A random variable x has pdf (probability density function)

$$f_x(x) = \begin{cases} 0.2 & -1 < x \leq 0 \\ -0.2 + 4x & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

a) Find c

b) Find $P(-0.5 < x < 0.5)$

$$(a) \int_{-\infty}^{\infty} f(x) dx = 1$$

$$\int_{-\infty}^{-1} f(x) dx + \int_{-1}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^{\infty} f(x) dx = 1$$

$\underbrace{\hspace{1cm}}$ $\underbrace{\hspace{1cm}}$ $\underbrace{\hspace{1cm}}$ $\underbrace{\hspace{1cm}}$

$$\int_{-1}^0 0.2 dx + \int_0^1 0.2 + cx dx$$

$$= [0.2x]_{-1}^0 + [0.2x + \frac{cx^2}{2}]_0^1 = 1$$

$$= (0 - (-0.2)) + (0.2 + \frac{c}{2} - 0 - 0) = 1$$

$$\frac{c}{2} = 0.6$$

$$c = 1.2$$

$$(b) P(-0.5 < X < 0.5) = \int_{-0.5}^{0.5} f(x) dx$$

$$= \int_{-0.5}^0 f(x) dx + \int_0^{0.5} f(x) dx$$

$$= [0.2x]_0^0 + [0.2x + \frac{1.2x^2}{2}]_0^{0.5}$$

$$= 0.1 + [0.1 + 0.6 \times 0.5^2]$$

Important absolute continuous distribution

① Uniform distribution

A random variable X is said to have a uniform distribution over the interval $[L, R]$ ($X \sim \text{uniform}[L, R]$) if its PDF is given

$$f_x(x) = \begin{cases} \frac{1}{R-L}, & L \leq x \leq R \\ 0, & \text{otherwise} \end{cases}$$

eygix1inquiry6lgdk7pgedl

$$\text{Note: } \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^L f(x) dx + \int_L^R f(x) dx + \int_R^{\infty} f(x) dx$$

$$= \left[\frac{1}{R-L} x \right]_L^R = \frac{1}{R-L} [x]_L^R = \frac{1}{R-L} (R-L) = 1$$

Ex/ $X \sim \text{Uniform} [-5, 6]$

Find (a) $P(1 < X < 3)$

find density function

$$f_x(x) = \frac{1}{6-5} = \frac{1}{1} = 0 \quad -5 \leq x \leq 6$$

0/w before -5 = 0

$$P(1 < x < 3) = \int_1^3 f(x) dx = \int_1^3 \frac{1}{1} dx = \frac{1}{1} x^3 \Big|_1^3 = \frac{1}{1} (3-1) = \frac{2}{1}$$

b) $P(x < 1)$

815570290

$$\begin{aligned} &= \int_{-\infty}^{-1} f(x) dx = \int_{-\infty}^{-5} f(x) dx + \int_{-5}^{-1} f(x) dx = \int_{-5}^{-1} \frac{1}{1} dx = \frac{1}{1} (x) \Big|_{-5}^{-1} = \frac{4}{1} \end{aligned}$$

$$c) P(|x| \leq 0.25) = P(-0.25 \leq x \leq 0.25)$$

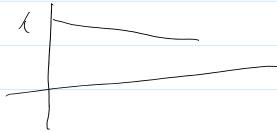
$|x| < x \Leftrightarrow -a < x < a$

$$\begin{aligned} &= \int_{-0.25}^{0.25} f(x) dx \\ &= \lambda [x] \Big|_{-0.25}^{0.25} = \frac{0.5}{1} \end{aligned}$$

② Experimental Distribution:

A random variable x is said to have an exponential distribution with parameter $\lambda (> 0)$ if its p.d.f. is

$$f_x(x) = \lambda e^{-\lambda x} = 0 \quad , \quad \lambda > 0$$



Note: is $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$ a p.d.f.?

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^{\infty} f(x) dx = \int_0^{\infty} \lambda e^{-\lambda x} dx = -[-e^{-\lambda x}] \Big|_0^{\infty} = 1(0) = 1$$

Ex/ $x \sim \text{exponential } (0^2)$

find $P(x > 10)$

$$\text{Ans} \quad \int_0^{\infty} f(x) dx$$

$$= \int_0^{\infty} f(x) dx = \int_0^{\infty} 0.2 e^{-0.2x} dx$$

$$= \left[-e^{-0.2x} \right]_0^{\infty}$$

$$= e^{0.2 \times 0}$$

$$= e^{-2}$$

$$P_x(x) = 0.2 e^{-0.2x} = 0.2 e^{-0.2x}$$

STAB52 Page 29

June 11, 2018 9:12 PM

Term Test: July 6th, 19:00 - 21:00

Room A4112, A4223 → Allowed chalkstart
both sides

Last hr * Continuous random variables

A random variable is continuous if $P(x=x) = 0 \forall x \in \mathbb{R}$

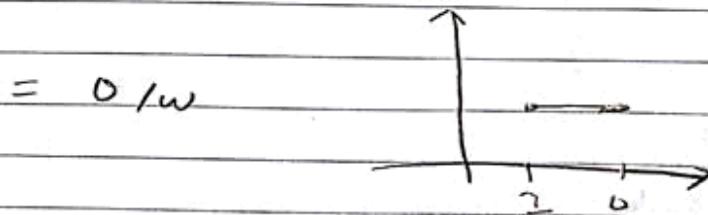
* Abs continuous random variables

A random variable is absolute continuous if \exists a p.d.f f_x s.t $P(a < x \leq b) = \int_a^b f_x(x) dx$

Important continuous distributions:-

① A uniform distribution on $[L, R]$

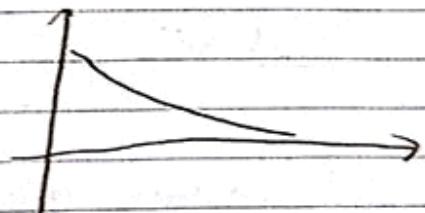
A random variable is said to have a uniform distribution on $[L, R]$ if $f_x(x) = \frac{1}{R-L}$ if $L \leq x \leq R$



② Exponential distribution

A random variable x is said to have an exponential distribution with the parameter $\lambda > 0$

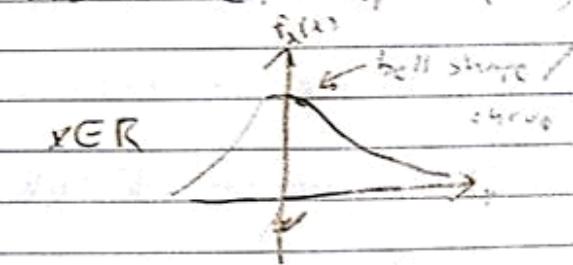
$$f_x(x) = \lambda e^{-\lambda x} \quad \text{if } x \geq 0 \\ = 0 \quad \text{otherwise}$$



To day: Normal distribution

A random variable x is said to have a normal distribution with parameters μ and σ^2 ($\mu \in \mathbb{R}$, $\sigma > 0$) if:

$$f_x(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \quad x \in \mathbb{R}$$



when $\mu=0$ and $\sigma=1$ is a special case

If normal distribution with $\mu=0$ and $\sigma=1$ is called a Standard Normal Distribution

pdf of standard normal distribution: $F_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$

PPF condition: $f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} > 0 \quad \forall x, x \in \mathbb{R}$

- PDF 2 conditions $\frac{L}{\int f(r) dr} \geq 0$

Condition ①:

$$\iint f(x, y) dx dy = \iint f(r \cos \theta, r \sin \theta) r dr d\theta \text{ Recall, integral across volume}$$

$$[] \int_a^b F(x) dx$$

$$Q^2, \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1$$

$$\Rightarrow \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}$$

$$L_1 + I = \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} dx \quad \text{and} \quad I = \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy$$

$$I^2 = \left[\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx \right] \left[\int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy \right]$$

$$= \iint_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2 - \frac{1}{2}y^2} dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x^2+y^2)} dx dy$$

$$= \iint_0^{2\pi} e^{-\frac{1}{2}(r^2 \cos^2 \theta + r^2 \sin^2 \theta)} r dr d\theta$$



$$= \iint_0^{2\pi} e^{-\frac{1}{2}r^2} r dr d\theta$$

$$\therefore I^2 = 2\pi$$

$$= \int_0^{2\pi} \left[e^{-\frac{1}{2}r^2} \right]_0^{2\pi} d\theta$$

which means
 $I = \sqrt{2\pi}$

$$= \int_0^{2\pi} 1 d\theta = [\theta]_0^{2\pi} = 2\pi$$

Term Test: July 6th 1:00-21:00 (1 hr 50 min)

psya

Rooms: AA112, AC223

Last hour: Important absolute continuous distributions

① Uniform $[L, R]$ distributions ($L \leq R$)

$$f_x(x) = \begin{cases} \frac{1}{R-L} & L \leq x \leq R \\ 0 & \text{o/w} \end{cases}$$

② Exponential Distributions ($x \geq 0$)

$$f_x(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{o/w} \end{cases}$$

③ Normal distribution ($x \sim N(\mu, \sigma^2)$)

$$f_x(x) = \frac{1}{\sigma \sqrt{\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$$

$N(\mu=0, \sigma^2=1)$ is called the standard normal distribution.

④ Gamma ($\sim \Gamma(\alpha)$) distribution ($x \geq 0, \alpha > 0$)

Recall $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \leftarrow \text{gamma function}$

Some properties includes:

① $\Gamma(1) = 1$

② if $\alpha > 1$, then $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

③ if α positive integer $\Gamma(\alpha) = (\alpha-1)!$

④ $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$$\begin{aligned} \text{① } \Gamma(1) &= \int_0^\infty t^{1-1} e^{-t} dt \\ &= \int_0^\infty e^{-t} dt \\ &= \left[-e^{-t} \right]_0^\infty \\ &= 1 \end{aligned}$$

Gamma distribution

we know $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$

so,

$$\int_0^\infty t^{\alpha-1} e^{-tx} dt, \text{ sub } t = ux, u > 0$$

$$\int_0^\infty t^{\alpha-1} e^{-tx} dt = \lambda x, \quad x > 0$$

(sub $t = \lambda x$, $\lambda > 0$)

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

$$\frac{\Gamma(\alpha)}{\Gamma(\alpha)} = \int_0^\infty \underbrace{\lambda^{\alpha-1} x^{\alpha-1} e^{-\lambda x}}_{\Gamma(\alpha)} dx = + \geq 0$$

\uparrow

$$f_x(x) = \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$

is a PDF

A random variable x is said to have a gamma distribution with parameters α and λ ($\alpha > 0, \lambda > 0$) if $f_x(x) = \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$ $x \geq 0$
 $= 0$ otherwise
 $(x \sim \text{Gamma}(\alpha, \lambda))$

Important Note:

Gamma ($\alpha=1, \lambda$)

$$f_x(x) = \frac{\lambda^x x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} = \frac{\lambda^x x^{\alpha-1} e^{-\lambda x}}{\Gamma(1)} = \lambda^x e^{-\lambda x}$$

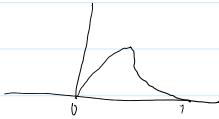
Note Gamma (λ) = exponential λ

⑤ Beta Distribution

A random variable x is said to have a beta distribution with parameters ($A > 0, B > 0$)

$$\text{if } F_x(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}, \quad 0 \leq x \leq 1$$

$$= 0 \quad 0 / \omega$$



$$\int_0^1 \frac{\lambda^{\alpha-1} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} dx = 1 \Rightarrow \int_0^1 x^{\alpha-1} e^{-\lambda x} dx = \frac{\Gamma(\alpha)}{\lambda^\alpha} \Big/ \int_0^1 x^{\alpha-1} e^{-\lambda x} dx / \alpha$$

$$= \frac{\Gamma(\alpha+1)}{\lambda^\alpha}$$

$$\int_0^1 x^a (1-x)^b dx$$

=

$$F_X(x) = \int \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} dx$$

$$= 0$$

or

$$= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

2.5 Cumulative distribution function (cdf)

Definition (cdf): Given a random variable x , its cumulative distribution function is the function $F_x: \mathbb{R} \rightarrow [0, 1]$ defined by $F_x(x) = P(x \leq x)$

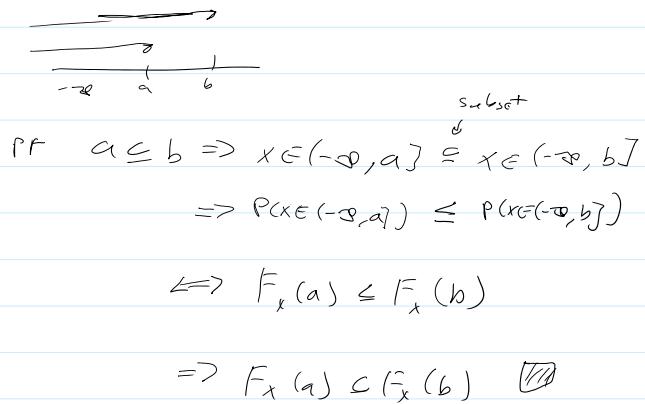
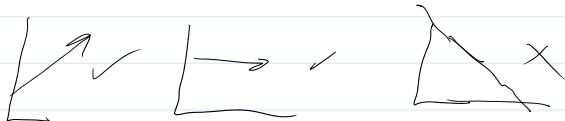
$$P(x \leq x) = P(x \in [-\infty, x])$$

Properties of cdf (F_x)

- ① $0 \leq F_x(x) \leq 1 \quad \forall x \in \mathbb{R}$
- ② $a < b \Rightarrow F_x(a) \leq F_x(b)$

If $a < b \Rightarrow f(a) \leq f(b)$

proves $f(x)$ is a non decreasing function



$$\textcircled{3} \lim_{x \rightarrow \infty} F_x(x) \approx 1$$

$$F_x(x) = P(x \leq x)$$

$$\textcircled{4} \lim_{x \rightarrow -\infty} F_x(x) = 0$$

$$\textcircled{5} \text{ for } a < b, P(a < x \leq b) = F_x(b) - F_x(a)$$



$$P(a < x \leq b) = P(x \in (a, b])$$

$$x \in (a, b] = x \in (-\infty, b) \setminus x \in (-\infty, a]$$

$$P(x \in (a, b]) = P(x \in (-\infty, b] \setminus x \in [a, \infty)) \quad , \text{ Rule of } P(A \setminus B) = P(A) - P(AB)$$

$$= P(x \in (-\infty, b]) - P(x \in (-\infty, a]) \dots (x \in (-\infty, a) \cup x \in (a, b]) \quad \text{if } B \subset A \\ A \setminus B = B$$

$$= P(x \leq b) - P(x \leq a)$$

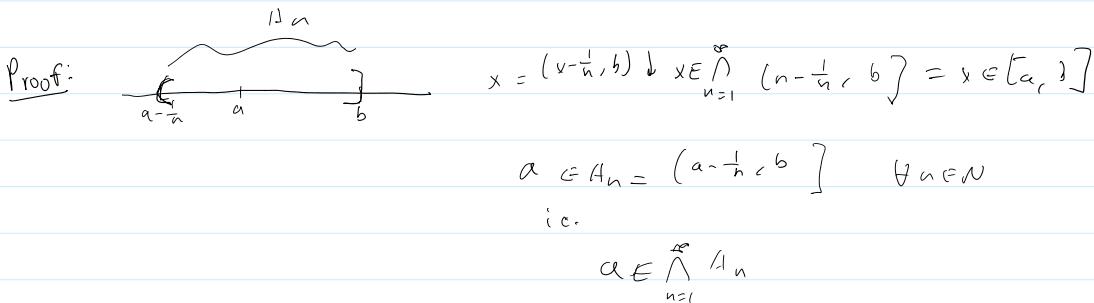
$$= F(b) - F(a)$$



Property 6

$$\textcircled{6} \quad P(a \leq x \leq b) = P(x \in [a, b]) = F_x(b) - F_x(a^-)$$

where $F_x(a^-) = \lim_{n \rightarrow \infty} F_x(a - \frac{1}{n})$



Using continuity of probability:

$$\begin{aligned} P(x \in [a, b]) &= \lim_{n \rightarrow \infty} P(x \in (a - \frac{1}{n}, b]) \\ &= \lim_{n \rightarrow \infty} [F_x(b) - F_x(a - \frac{1}{n})] \\ &= \lim_{n \rightarrow \infty} F_x(b) - \lim_{n \rightarrow \infty} F_x(a - \frac{1}{n}) \\ &= F_x(b) - F_x(a^-) \quad \blacksquare \end{aligned}$$

$$\textcircled{7} \quad P(x \in (a, b)) = F_x(b^-) - F_x(a) \quad (\text{Exercise prove it})$$

$$\textcircled{8} \quad P(x \in [a, b]) = F_x(b^-) - F_x(a^-)$$

CDF's of discrete random variables

Ex Bernoulli(θ)

$$\boxed{x | p_x} \quad x \sim \stackrel{i.i.d}{\sim} \text{Ber}(\theta)$$

x	P _x
0	1-θ
1	θ

$$x \sim \text{Ber}(\theta)$$

$$F_x(-0.5) = P(x \leq -0.5)$$

$$= 0$$

← since $\text{Ber}(\theta)$ can only be 0, 1



$$F_x(0) = P(x \leq 0) = P(x=0) = 1-\theta$$

$$F_x(0.5) = P(x \leq 0.5) = P(x=0) = 1-\theta$$

$$F_x(0.7) = 1-\theta$$

$$F_x(1) = P(x \leq 1)$$

$$= P(x=0) + P(x=1)$$

$$= 1-\theta + \theta$$

$$F_x(1.2) = P(x \leq 1.2)$$

$$= 1$$

Ex Let x be a random variable with pmf $p_x(x) = \frac{1}{4}$, $x = 1, 2, 3, 4$

Plot the graph of $F_x(x)$

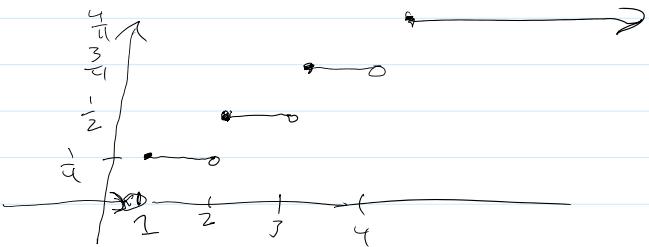
$$F_x(0.5) = P(x < 0.5)$$

$$= 0$$

$$F_x(1) = P(x \leq 1)$$

$$F_x(2) = P(x \leq 2) = P(x=2) + P(x=1)$$

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

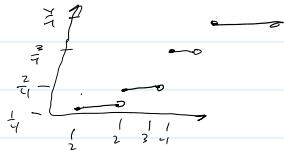


Term test: July 6 . 19:00-21:00 (1 hr 50 min)

Rooms AA112, AC223

Last-hour: cdf's

$$\text{Ex } P_x(x) = \begin{cases} 1 & x \in \{1, 2, 3, 4\} \\ 0 & \text{o/w} \end{cases}$$



$$F_x(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{4}, & 1 \leq x < 2 \\ \frac{2}{4}, & 2 \leq x < 3 \\ \frac{3}{4}, & 3 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$

$$F_x(0.5) = 0$$

$$F_x(2) = \frac{2}{4} = \frac{1}{2}$$

$$F_x(2^-) = \lim_{n \rightarrow \infty} F_x(2 - \frac{1}{n}) = \frac{1}{4} \neq F_x(2)$$

$$F_x(2^+) = \lim_{n \rightarrow \infty} F_x(2 + \frac{1}{n}) = \frac{2}{4} = \frac{1}{2} = F_x(2)$$

} $f(x)$ is continuous
 $a^- = a$ if
 $\lim_{x \rightarrow a^-} f(x) = f(a)$
 $\lim_{x \rightarrow a^-} f(x) = f(a)$, left continuous

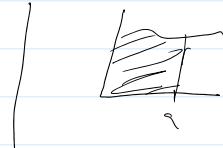
F_x is right continuous at 2 but F_x is not left continuous at 2

$\lim_{x \rightarrow a^+} f(x) = f(a)$, right continuous

Note: The cdf's are right continuous and \Rightarrow [cadlag functions]
 have finite left limits.

CDF's of continuous random variables

$$F_x(x) = P(X \leq x) = \int_{-\infty}^x f_x(t) dt$$



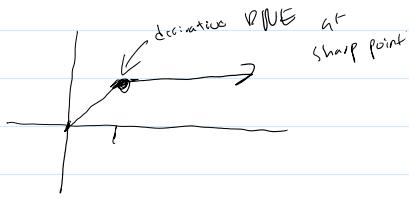
$$f_x(x) = \frac{d}{dx} F_x(x)$$

Ex A r.v x has cdf given by $F_x(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$

Find the pdf of x .

W.M.I.S. .+

Find the pdf of x ,



$\sim \mathcal{N}(0, \sigma^2)$

$$f_x(x) = \frac{d}{dx} F_x(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases}$$

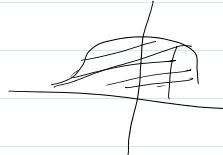
← pay attention
← us ←
be careful.

CDF of the standard normal distribution ($N(0, \sigma^2 = 1)$)

$$\Phi(x) = P(X \leq x) = \int_{-\infty}^x \phi(t) dt$$

$$= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

← only calc calc help



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

Ex $X \sim N(0, 1)$, find $P(-0.63 < X < 2.00)$

For many cont $\subset = \subseteq$

$$P(X \in (a, b)) = P(a < X \leq b) = F_x(b) - F_x(a)$$

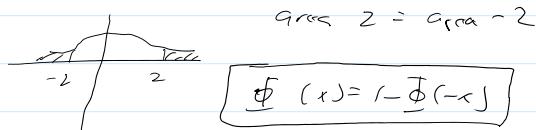
≈ 0.2643

$$= \Phi(2.00) - \Phi(-0.63)$$



Table on pg 712

Z	0.0	-0.1	-0.2	...	0.9
-3.4					
-3.3					
-3.2					
-3.1					
-3.0					
-2.9					
-2.8					
-2.7					
-2.6					
-2.5					
-2.4					
-2.3					
-2.2					
-2.1					
-2.0					
-1.9					
-1.8					
-1.7					
-1.6					
-1.5					
-1.4					
-1.3					
-1.2					
-1.1					
-1.0					
-0.9					
-0.8					
-0.7					
-0.6					
-0.5					
-0.4					
-0.3					
-0.2					
-0.1					
0.0					
0.1					
0.2					
0.3					
0.4					
0.5					
0.6					
0.7					
0.8					
0.9					
1.0					
1.1					
1.2					
1.3					
1.4					
1.5					
1.6					
1.7					
1.8					
1.9					
2.0					
2.1					
2.2					
2.3					
2.4					
2.5					
2.6					
2.7					
2.8					
2.9					
3.0					
3.1					
3.2					
3.3					
3.4					



$$= \Phi(2.00) - \Phi(-0.63)$$

$$= (1 - \Phi(-2.00)) - 0.2643$$

$$= (1 - 0.228) - 0.2643$$

Probabilities for the $N(\mu, \sigma^2)$

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} dx$$

, substitute $z = \frac{x-\mu}{\sigma}$

$$= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} z^2} dz = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$dz = \frac{1}{\sigma} dx$$

$$\begin{aligned}
 &= \int_{\frac{\alpha-\mu}{\sigma}}^{\frac{\beta-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
 &\Rightarrow \left[\Phi(z) \right]_{\frac{\alpha-\mu}{\sigma}}^{\frac{\beta-\mu}{\sigma}} \\
 &= \Phi\left(\frac{\beta-\mu}{\sigma}\right) - \Phi\left(\frac{\alpha-\mu}{\sigma}\right)
 \end{aligned}$$

if $x \sim N(\mu, \sigma^2)$ then $P(a < x < b) = \boxed{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}$

Ex/ (Ex 2.5.5)

$$y \sim N(-8, 4), \text{ Find } (a) P(-2 < y \leq 7) \quad (b) P(y \geq 3)$$

$$\begin{aligned}
 a) P(-2 < y \leq 7) &= \Phi\left(\frac{7-(-8)}{2}\right) - \Phi\left(\frac{-2-(-8)}{2}\right) \\
 &= \Phi(7.5) - \Phi(3) \\
 &= 0.0013
 \end{aligned}$$

$$\begin{aligned}
 b) P(y > 3) &= 1 - P(y \leq 3) = 1 - \Phi\left(\frac{3-\mu}{\sigma}\right) \\
 &= \Phi(\infty) - \Phi\left(\frac{3-(-8)}{2}\right) \\
 &= 1 - \Phi(5.5) \\
 &= 1 - (1 - \Phi(-5.5)) \\
 &= 0
 \end{aligned}$$

Midterm Test July 6th 19:00-21:00 (1 hr 50 min)

Rooms A112, A223

Last hour: Function of random variables

Question: X has a pdf f_X and $y = h(x)$ we want f_Y

Two methods.

① CDF method:

Find $F_Y(y)$ using $F_Y(y) = P(Y \leq y) = P(h(X) \leq y)$ ← method has no conditions.

and then $f_Y(y) = \frac{d}{dy} F_Y(y)$

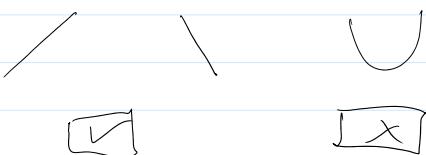
Today (Method 2: One dimensional transformation formula)

Midterm upto and including 2.5. 2.6 is not on midterm.

7 questions, each question is 10 points. You can use both sides.

Theorem 2.6.3

Let X be an absolute continuous random variable and $y = h(x)$ where $h: \mathbb{R} \rightarrow \mathbb{R}$ is differentiable and strictly increasing OR strictly decreasing on the support of f_X



Then y is absolute continuous and $f_Y(y) = \frac{f_X(h^{-1}(y))}{|h'(h^{-1}(y))|}$ ← not derivative of composite function

Where $h'(x) = \frac{d}{dx} h(x)$

Support (f) = { $x : f(x) > 0$ }

Example: If random variable X has pdf

$$f_X(x) = \begin{cases} 2(1-x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

L o O/w

Find the pdf of

a) $y = 2x - 1 = h(x)$

check if method requirements are suitable

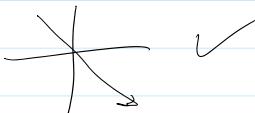
① $y = 2x - 1$  straight line differentiable everywhere ✓

② $h^{-1}(y) = \frac{1+y}{2}$ ✓

③ $h'(x) = \frac{d h(x)}{dx} = 2$

$$f_y(y) = \frac{f_x(h^{-1}(y))}{|h'(h^{-1}(y))|} = \frac{2\left(1 - \frac{1+y}{2}\right)}{|2|}, \quad 0 \leq \frac{1+y}{2} \leq 1$$
$$= \left(\frac{1-y}{2}\right), \quad -1 \leq y \leq +1$$
$$= 0 \quad \text{o/w}$$

b) $y = 1-2x = h(x)$

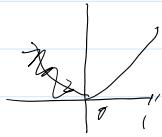
①  ✓

② $h^{-1}(y) = \frac{1-y}{2}$

③ $h'(x) = \frac{d h(x)}{dx} = -2$

$$f_y(y) = \frac{f_x(h^{-1}(y))}{|h'(h^{-1}(y))|} = \frac{2\left(1 - \frac{1-y}{2}\right)}{|-2|}, \quad 0 \leq \frac{1-y}{2} \leq 1$$
$$= \left(1 - \frac{1-y}{2}\right) \quad / \quad 0 \leq \frac{1-y}{2} \leq 1$$
$$= 0 \quad \text{o/w}$$

c) $y = x^2$, you can use it since it's increasing from $(0, 1)$



$$y = x^2 = h(x), \quad h^{-1}(y) = \sqrt{y}, \quad h'(x) = \frac{d h(x)}{dx} = 2x$$

$$f_y(y) = f_x(h(y)) = \frac{2(1-\sqrt{y})}{2\sqrt{y}}, \quad 0 \leq y \leq 1$$

$$= \frac{1-\sqrt{y}}{\sqrt{y}}, \quad 0 < y \leq 1$$

$$h'(h^{-1}(y)) = 2\sqrt{y}$$

$$= 0, \quad \text{otherwise}$$

2.7 joint distributions

Definition:

distribution of random variable
 $P(x \in B), \quad \forall B \subseteq \mathbb{R}$

The joint distribution of two random variables x and y is the collection of probability

$$P((x,y) \in B), \quad \forall B \subseteq \mathbb{R}^2$$

$$(x,y) \in B = \{s: x(s), y(s) \in B\}$$

Definition joint cdf:

The joint cdf of random variables x & y is the function $F_{x,y}: \mathbb{R}^2 \rightarrow [0,1]$ defined by

$$F_{x,y}(x,y) = P(x \leq x, y \leq y)$$

recall,
cdf $F_x(x) = P(x \leq x)$
 $F_x: \mathbb{R} \rightarrow [0,1]$

Marginal distribution

$$F_x(x) = \lim_{y \rightarrow \infty} F_{x,y}(x,y) = F_{x,y}(x, \infty)$$

$$F_y(y) = \lim_{x \rightarrow \infty} F_{x,y}(x,y)$$

" " = \wedge = and

- Notes
- ① $F_{x,y}(\infty, \infty) = 1$
 - ② $F_{x,y}(-\infty, -\infty) = 0$
 - ③ $F_{x,y}(-\infty, y) = 0$
 - ④ $F_{x,y}(x, -\infty) = 0$

Definition Joint Probability mass function

What is PMF: only for discrete random variables
 $P_x(x) = P(x=x)$

Let x and y be discrete random variables. The joint pmf of x & y is the function P_{xy} : $\mathbb{R}^2 \rightarrow [0, 1]$ defined by $P_{xy}(x, y) = P(x=x, y=y)$

Consider an experiment of tossing a fair coin and a six sided die. $x =$ outcome of the die, $y =$ flip on the coin (0 or 1). Find the joint pmf of x and y

	x	1	2	3	4	5	6
y	0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$
	1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

$$P_x(x) = e^{-\lambda} \frac{\lambda^x}{x!}, x=0, 1, \dots$$

= 0 otherwise



$$S = \{(1,0), (1,1), (2,0), (2,1), (3,0), (3,1), (4,0), (4,1), (5,0), (5,1), (6,0), (6,1)\}$$

Term Test: July 6th, 19:00-21:00 (1 hr 50 min) (1.0 - 2.5)

Last hour Section 2.7 joint distributions.

For discrete random variables $x=y$, the joint pmf

$$P_{x,y}(x,y) = P(x=x, y=y)$$

$$\text{Marginal PDF of } x: P_x(x) = \sum_y P_{x,y}(x,y) \quad | \quad P_y(y) = \sum_x P_{x,y}(x,y)$$

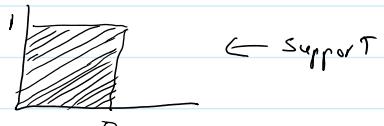
Joint density function satisfies

$$\textcircled{1} \quad f_{x,y}(x,y) \geq 0 \quad \textcircled{2} \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{x,y}(x,y) dx dy = 1$$

Random variables x and y are jointly absolutely continuous if $\exists f_{x,y}$ s.t. $P(a < x \leq b, c < y \leq d) = \int_a^b \int_c^d f_{x,y}(x,y) dx dy$

Ex (done)

$$f_{x,y}(x,y) = \begin{cases} Kxy, & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$



(a) Find K ($A_{\text{support}} K = 1$)

(b) Find $P(x \leq \frac{1}{2}, y \leq \frac{3}{4})$

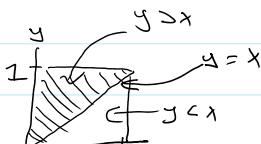
Next week Quiz 2 questions: One based 2.6 and one 2.7

Today

Ex Random variables $x+y$ have joint pdf

$$f_{x,y}(x,y) = \begin{cases} K(x-y) & 0 \leq x \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find K
 $\Rightarrow \pi$



Find K

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

$$= \int_0^1 \int_0^y K(1-y) dx dy$$

$$= K \int_0^1 \int_0^y (1-y) dx dy$$

$$= K \int_0^1 (1-y) \int_0^y dx dy$$

$$= K \int_0^1 (1-y) [x]_0^y dy$$

$$= K \int_0^1 (y - y^2) dy$$

$$> K \left[\frac{y^2}{2} - \frac{y^3}{3} \right]_0^1$$

$$1 = \frac{1}{6} K \Rightarrow K = 6$$

Part B

common = \cap = intersection

Find the $P(x \leq \frac{3}{4}, y \geq \frac{1}{2})$

$$= \int_{\frac{1}{2}}^{\frac{3}{4}} \int_0^1 6(1-y) dx dy + \int_{\frac{3}{4}}^1 \int_0^1 6(1-y) dx dy$$

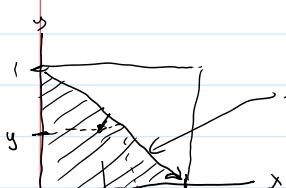
$$= 6 \int_{\frac{1}{2}}^{\frac{3}{4}} (1-y) \underbrace{[x]_0^y dy}_{(y^2 - y^3)} + 6 \int_{\frac{3}{4}}^1 (1-y) \underbrace{[x]_0^y dy}_{(y^2 - y^3)}$$

$$= 6 \left[\frac{y^2}{2} - \frac{y^3}{3} \right]_{\frac{1}{2}}^{\frac{3}{4}} + 6 \times \frac{3}{4} \left[y - \frac{y^2}{2} \right]_{\frac{3}{4}}^1$$

= ? Ex

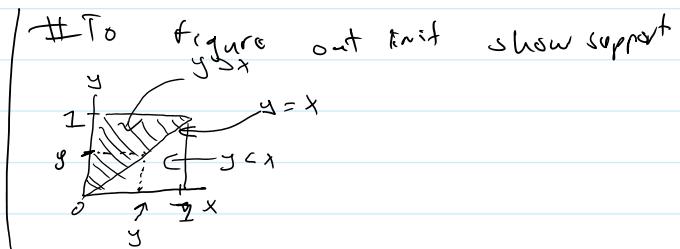
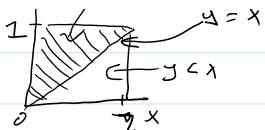
$$\text{Ex } f_{xy}(x,y) = \begin{cases} 4xy & , 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & , \text{ otherwise} \end{cases}$$

Find $P(x+y \leq 1)$



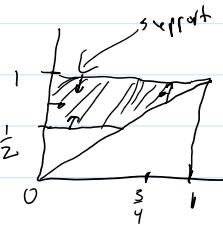
$$\begin{aligned} P(x+y \leq 1) &= \int_0^1 \int_0^{1-x} 4xy dx dy \\ &= 4 \int_0^1 y \left[\frac{x^2}{2} \right]_0^{1-y} dy \end{aligned}$$

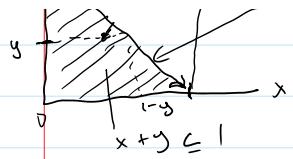
$$\begin{aligned} &\text{R.C.C.} \\ &\int_0^1 x^{a-1} (1-x)^{b-1} dx \end{aligned}$$



Ex / integrate the integral
by changing $dy dx$

.





$$x+y=1$$

then

$$x = 1-y$$

$$= \int_0^1 y \left[\frac{y^2}{2} \right]_0^1 dy$$

$$= \frac{1}{2} \int_0^1 y(1-y^2) dy$$

$$= 2 \times \frac{\Gamma(2)\Gamma(3)}{\Gamma(5)} =$$

$$= \frac{2x+x^2}{4! \cdot 2 \cdot 4} \times 2 =$$

$$\int_0^1 x^{a-1} (1-x)^{b-1} dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Final exam type question

Ex Random variable x and y has joint density

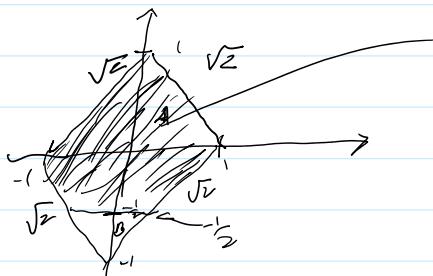
$$f_{xy}(x,y) = \begin{cases} K, & \text{if } |x| + |y| \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\text{Find } P(y \geq -\frac{1}{2})$$

you don't need to find K .

$$|x| + |y| \leq 1$$

$$1 - \frac{1}{2} \left(\frac{1}{2} \right) \times 1$$



Any point is equally likely to receive the point

Marginal distributions

$$\text{PMF} = P_{xy}(x,y) = P(x=x, y=y)$$

$$P_x(x) = \sum_y P_{xy}(x,y)$$

$$P_y(y) = \sum_x P_{xy}(x,y)$$

^{then} If x and y are jointly absolute continuous random variables with joint pdf $f_{xy}(x,y)$ then the marginal pdf's are given by

$$f_x(x) = \int_{-\infty}^{\infty} f_{xy}(x,y) dy, \quad \text{if } f_x(y) = \int_{-\infty}^{\infty} f_{xy}(x,y) dx$$

$$\text{Ex } f_{xy}(x,y) = \begin{cases} 4xy & , 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & , \text{ otherwise} \end{cases}$$

Find f_x : marginal pdf of f_x

Answer

$$f_x(x) = \int_0^1 f_{xy}(x,y) dy = \int_0^1 4xy dy = 4x \int_0^1 y dy = 4x \left[\frac{y^2}{2} \right]_0^1$$

$$= 2x, \quad 0 \leq x \leq 1$$

$$\text{Ex } F_y(y) ? \quad (\text{ans } 2y), \quad 0 \leq y \leq 1$$

$$= 0$$

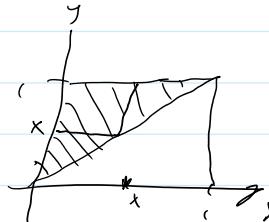
remember
this

$$\text{Ex } f_{xy}(x,y) = \begin{cases} 6(1-y) & , 0 \leq x \leq y \leq 1 \\ 0 & , \text{ otherwise} \end{cases}$$

Find f_x

$$\text{Ans } f_x(x) = \int_x^1 6(1-y) dy$$

$$= 6 \left[\frac{(1-y)^2}{2} \right]_x^1$$



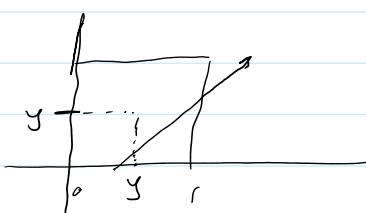
$$= 3(1-x)^2, \quad 0 < x < 1$$

$$= 0, \quad \text{otherwise}$$

b) Find $f_y(y)$

$$f_y(y) = \int 6(1-y) dx$$

$$= \int_0^y 6(1-y) dx$$



$$\begin{aligned}
 &= 6(1-y) [x \times f_y] \\
 &= 6y(1-y) \quad 0 < y < 1 \\
 &= 0 \quad \text{otherwise}
 \end{aligned}$$

2.8 Conditional distribution

Definition Let x and y be jointly discrete random variables then for any x s.t $p(x=x) > 0$, the conditional distribution of y given $x=x$ is the collection of probabilities.

$$\frac{P(y \in B, x=x)}{p(x=x)}, \quad \forall B \subseteq \mathbb{R}$$

In particular, this assigns probability

$$\frac{P(a < y \leq b, x=x)}{p(x=x)}, \quad \text{for the event } a < y \leq b$$

Definition Conditional mass function

Let x and y be jointly discrete with joint mass function P_{xy} and let $x \in \mathbb{R}$ be s.t $p(x=x) > 0$ [i.e. $p_x(x) > 0$]

then the conditional mass function of y given $x=x$ is given by

$$p_{y|x}(y|x) = P_{xy}(x,y) = \frac{p(x=x, y=y)}{p_x(x)}$$

Example

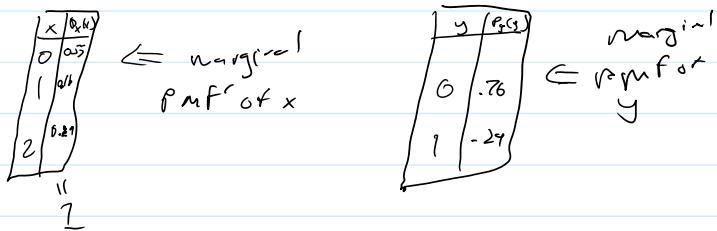
$$x = \begin{cases} 0, & \text{if no bolt used} \\ 1, & \text{if adult bolt used} \\ 2, & \text{if car seat bolt used} \end{cases}$$

$$y = \begin{cases} 0, & \text{if child survived} \\ 1, & \text{if not} \end{cases}$$

The joint mass function of x and y

$x \setminus y$	0	1
0	0.38	0.17
1	0.14	0.02
2	0.24	0.05
	.76	.24

a) Give the marginal pmf's of x and y



b) Give the conditional pmf mass function for y given $x=0$.

$$\text{Ans } P_{y|x}(y|0) = \frac{P_{xy}(0,y)}{P_x(0)}$$

y	$P_{y x}(y 0)$
0	$\frac{0.38}{0.55}$
1	$\frac{0.17}{0.55}$

$\frac{P_{xy}(0,1)}{P_x(0)}$

Definition (conditional density function)

Let x and y be jointly absolutely continuous with joint density function f_{xy} and let $x \in \mathbb{R}$ be s.t. $f_x(x) > 0$, then the conditional density function of y given $x=x$ is defined by $f_{y|x}(y|x) = f_{xy}($

Things to review:

5) Suppose that the random variable X has probability density function:

$$f(x) = \begin{cases} \frac{kx^2}{(1+x)^5} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where k is a constant.

Find the value of k that makes $f(x)$ a valid probability density function.

- A) 3
- B) 6
- C) 12
- D) 24

E) none of the values given in the choices (A) to (D) above, makes the function $f(x)$, a valid probability density function.

To find the constant k we integrate and set it equal to 1.

$$\int_0^\infty \frac{kx^2}{(1+x)^5} dx = 1$$

To evaluate this integral substitute $1+x = y^{-1}$, then $dx = (-1)y^{-2}dy$ and $x = \frac{1-y}{y}$ and the limits (0, infinity) changes to (1, 0).

$$\begin{aligned} k \int_0^\infty \frac{x^2}{(1+x)^5} dx &= k \int_1^0 \frac{(1-y)^2}{y^5} y^5 (-1)y^{-2} dy = k \int_0^1 y^1 (1-y)^2 dy = kB(2,3) = k \frac{\Gamma(2)\Gamma(3)}{\Gamma(5)} = k \frac{1 \times 2}{1 \times 2 \times 3 \times 4} \\ &= \frac{k}{12} \end{aligned}$$

Setting this to 1 we have $k = 12$

Discrete random variables

PDFS

CDFS

Continuity question

Standard Normal Table

Last hour: Section 2.8 conditional distribution

Definition if x and y are jointly discrete

$$P(Y \in B | x=x) = \frac{P(Y \in B, x=x)}{P(x=x)} \quad \text{for } P(x=x) > 0$$

Conditional probability

$$P_{y|x}(y|x) = \frac{P_{x,y}(x,y)}{P_x(x)}, \quad P_x(x) > 0$$

For x, y jointly absolutely continuous

$$\text{The conditional pdf: } f_{y|x}(y|x) = \frac{f_{x,y}(x,y)}{f_x(x)}, \quad f_x(x) > 0$$

$$\text{Definition } P(a \leq y \leq b | x) = \int_a^b f_{y|x}(y|x) dy$$

Independence of random variables

Definition: Two random variables x and y are independent if $P(x \in B_1, y \in B_2) = P(x \in B_1) \cdot P(y \in B_2) \quad \forall B_1, B_2 \subseteq \mathbb{R}$

if $A \& B$ are independent, we write $X \perp Y$

Thrm 2.8.2

$X \perp Y$ iff $P(a \leq x \leq b, c \leq y \leq d) = P(a \leq x \leq b) \cdot P(c \leq y \leq d)$
whenever $a \leq b, c \leq d$.

Thrm 2.8.3

a) if x and y are jointly discrete then $x \perp y$ iff

$$P_{X,Y}(x,y) = P_X(x) \cdot P_Y(y) \quad \forall x, y \in \mathbb{R}$$

b) if x & y are jointly absolutely continuous then x and y are \perp iff

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad \leftarrow \text{This one's easier! to prove } \perp$$

$$\text{Ex: } f_{X,Y}(x,y) = \begin{cases} 4xy & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Are x and $y \perp$?

$$f_X(x) = 2x \quad 0 \leq x \leq 1$$

$$f_Y(y) = 2y \quad 0 \leq y \leq 1$$

$$\begin{aligned} f_{X,Y}(x,y) &= 4xy = (2x)(2y) \\ &= f_X(x) f_Y(y) \end{aligned} \quad \therefore x \perp y \quad \boxed{\text{True}}$$

$$\text{Ex: } f_{X,Y}(x,y) = \begin{cases} 6(1-y) & , 0 \leq x \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Are x and $y \perp$?

$$f_X(x) = 3(1-x)^2, \quad 0 \leq x \leq 1 \quad \left| \quad x = \frac{1}{2}, y = \frac{1}{2} \right.$$

$$f_Y(y) = 6y(1-y), \quad 0 \leq y \leq 1 \quad \left| \quad f_X\left(\frac{1}{2}\right) = 6\left(1-\frac{1}{2}\right) = 3 \right.$$

$$\cdot f_X\left(\frac{1}{2}\right) \times f_Y\left(\frac{1}{2}\right) = 3\left[1-\frac{1}{2}\right]^2 \times 6 \times \frac{1}{2} \left(1-\frac{1}{2}\right) \neq 3$$

$x \perp y$ ← not independent

Thrm 2.8.4

a) if x and y are jointly discrete then

$x \perp y$ if $P_{y|x}(y|x) = P_y(y)$ $\forall x, y \in \mathbb{R}$ s.t. $P_x(x) > 0$

b) if x and y are jointly absolutely continuous, then

$x \perp y$ if $f_{y|x}(y|x) = f_y(y)$ $\forall x, y \in \mathbb{R}$ s.t. $f_x(x) > 0$

Chapter 3 Expectation (Mean)

Definition if x is a discrete random variable with pmf f_x , then the expectation of x (mean of x , expected value of x) is defined by

$$E(x) \text{ or } E X = \sum_x x p_x(x)$$

Ex/ If r.v x has pmf

x	$p_x(x)$
1	0.2
2	0.3
3	0.5

Find $E(x)$

$$\text{Ans} \quad E(x) = \sum_x x p_x(x) = (1 \times 0.2) + (2 \times 0.3) + (3 \times 0.5)$$

① $x \sim \text{degenerate at } c$, find $E(x)$?

Recall Degenerate distribution:

x	$p_x(x)$
c	1

$$\text{mean of } x = E(x) = c \times 1 = c$$

② $X \sim B_{\text{bin}}(\theta)$. Find $E(X)$

Ans:

x	$P_{X 0}$
0	$1-\theta$
1	θ

$$E(X) = \sum x P_X(x) = 0(1-\theta) + 1(\theta) = \theta$$

③ $X \sim \text{Bin}(n, \theta) = \text{Find } E(X)$

$$P_X(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, x=0, 1, \dots, n \quad \leftarrow \text{Binomial PMF}$$

$$\text{Ans: } E(X) = \sum_{x=0}^n x P_X(x) = \sum_{x=1}^n x P_X(x) \leftarrow \text{since the first time } x \text{ value is 0}$$

$$= \sum_{x=0}^n x \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

$$= \sum_{x=1}^n x \left(\frac{n!}{x!(n-x)!} \right) \theta^x (1-\theta)^{n-x}$$

$$= \sum_{x=1}^n \frac{n!}{(x-1)!(n-x)!}$$

$$= n \theta \sum_{x=1}^n \frac{(n-1)!}{(x-1)![n-(x-1)]!} \theta^{x-1} (1-\theta)^{[(n-1)-(x-1)]}$$

$$\text{let } y = x-1$$

recall $(a+b)^n$

$$= n \theta \sum_{y=0}^{n-1} \frac{(n-1)!}{(y)![n-(y)]!} \theta^y (1-\theta)^{(n-1)-y}$$

$$= \sum_{x=0}^n \binom{n}{x} a^x b^{n-x}$$

$$= \sum_{x=0}^n \frac{n!}{x!(n-x)!} a^x b^{n-x}$$

$$= n \theta (1+\theta)^{n-1}$$

$$= n\theta$$

If x is bin (n, θ) , then the $E(x) = n\theta$

Basic Math (Recall)

Suppose that a power series $\sum_{n=0}^{\infty} a_n (x-c)^n$ has

radius of convergence R , then $\sum_{n=0}^{\infty} a_n (x-c)^n$ converges so defines a function $f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n$

- $f(x)$ has derivatives of all orders.

$f'(x) = \sum_{n=1}^{\infty} a_n \cdot n (x-c)^{n-1}$	$\begin{aligned} & \frac{d}{dx} \sum_{n=0}^{\infty} a_n (x-c)^n \\ &= \sum_{n=0}^{\infty} \frac{d}{dx} a_n (x-c)^n \end{aligned}$
$f''(x) = \sum_{n=2}^{\infty} a_n n(n-1) (x-c)^{n-2}$	

etc...

② $x \sim Ge(\theta)$, find $E(x)$? $P_x(x) = (re)^x \theta$, $r = 1 - \theta$

Ans $E_x = \sum_{x=0}^{\infty} x P_x(x) = \sum_{x=0}^{\infty} x (1-\theta)^x (\theta)$

$$= \theta \sum_{x=0}^{\infty} x (1-\theta)^x , \quad \text{let } q = 1-\theta$$

$$= \theta \sum_{x=1}^{\infty} x q^x$$

$$= \theta q \sum_{x=1}^{\infty} x q^{x-1}$$

$$= \theta q \sum_{x=1}^{\infty} \frac{d}{dq} q^x$$

$$= \theta q \frac{d}{dq} \left(\sum_{x=1}^{\infty} q^x \right) = \frac{\alpha}{1-r}$$

geometric series = $q^1 + q^2 + q^3 \dots$

$$= \theta q \frac{d}{dq} \left(\sum_{x=1}^{\infty} q^x \right) = \frac{\alpha}{1-r}$$

$$= \theta q \frac{d}{dq} \left(\frac{q}{1-q} \right) = \frac{\alpha}{1-q}$$

$$= \theta q \frac{d}{dq} \left(-1 + \frac{1}{1-q} \right)$$

$$= \theta q [0 + (1-q)^2]$$

$$= \theta q \left(\frac{1}{1-q} \right)^2 \leftarrow \text{sub } \theta \text{ back in}$$

$$= \theta q \left(\frac{1}{\theta} \right)^2$$

$$= \theta (1-\theta) \left(\frac{1}{\theta} \right)$$

$$= \frac{(1-\theta)}{\theta}$$

$$\therefore x \sim \text{Geo}(\theta) \text{ then } E(x) = \frac{1-\theta}{\theta}$$

Expect refors to $E(x)$

Ex Toss a biased coin with $P(H) = 0.1$ repeatedly. How many tails do you expect before your first head?

$x = \# \text{ of tails before the first head}$

then

$$x \sim \text{Geo}(\theta = 0.1)$$

$$E(x) = \frac{1-\theta}{\theta} = \frac{1-0.1}{0.1} = \frac{0.9}{0.1} = 9$$

(5) $x \sim P_0(x)$. Find $E(x)$: Recall $P_x(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, 2, \dots$

$$\text{Ans } E(x) \stackrel{\text{def}}{=} \sum_{x=0}^{\infty} x P_x(x) = \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!}, \quad e^{-\lambda} \text{ is constant}$$

$$x = 0 \quad \overline{x!}$$

$$= e^{-\lambda} \sum_{x=1}^{\infty} x \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!}$$

$$= e^{-\lambda} \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} \left[\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \dots \right]$$

$$= \lambda e^{-\lambda} e^{\lambda}$$

$$= \lambda \quad x \sim P_0(\lambda) \Rightarrow E(x) = \lambda$$

remember λ is
mean.

Ex A random variable x has

$$\text{pmf } p_x(x) = \frac{1}{x(x+1)}, \quad x = 1, 2, 3, \dots$$

Find $E(x)$

$$\text{Ans} \quad E(x) = \sum_x x p_x(x) = \sum_{x=1}^{\infty} x p_x(x)$$

$$= \sum_{x=1}^{\infty} x \cdot \frac{1}{x(x+1)} = \sum_{x=1}^{\infty} \frac{1}{x+1}$$

$$= \infty$$

tx show that a
pmf i.e.
 ① $p_x(x) \geq 0, \forall x$
 ② $\sum_{n=1}^{\infty} \frac{1}{x(x+1)} = 1$

expected to know

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

Ex A r.v x has pmf

$$p_x(x) = \begin{cases} \frac{1}{2^n} & x = 2, 4, 8, 16, \dots \\ \frac{1}{2} & x = -2, -4, -8, -16, \dots \end{cases}$$

Find $E(x)$

$$E(x) = \sum_{x} x p_x(x) = \sum_{n=1}^{\infty} 2^n \frac{1}{2 \cdot 2^n} + \sum_{k=1}^{\infty} (-2^k) \frac{1}{2 \cdot 2^n}$$

$$= \sum_{n=1}^{\infty} 2^n \frac{1}{2 \cdot 2^n} + (-1) \sum_{n=1}^{\infty} (2^n) \frac{1}{2 \cdot 2^n}$$
$$= \infty - \infty$$
$$= \text{D}$$

$E(y)$ is undefined, when its ∞ , it is defined.

Last hour Chapter 3 Expectation / mean / expected value

- * Discrete random variables [1]
- * Definition $E(X) = \sum_x P_x(x)$
- * $X \sim \text{Degenerate}(c) \Rightarrow E(X) = c$
- * $X \sim \text{Ber}(\theta) \Rightarrow E(X) = \theta$
- * $X \sim \text{Bin}(n, \theta) \Rightarrow E(X) = n\theta$
- * $X \sim \text{Geo}(\theta) \Rightarrow E(X) = \frac{1-\theta}{\theta}$
- * $X \sim P_0(\lambda) \Rightarrow E(X) = \lambda$
- * $X \sim f_x(x) = \frac{1}{x(\text{rel})}, x=1, 2, \dots \Rightarrow E(X) = \infty$
- * $X \sim f_x(x) = \begin{cases} \frac{1}{2^x}, & x=2^4, 8, \dots, 2^n \\ \frac{1}{2^{(x+1)}}, & x=-2, -4, -8, \dots, -2^n \end{cases} \quad \left. \begin{array}{l} \Rightarrow E(X) \text{ does not exist} \\ \text{if } E(X) = \infty - \infty \end{array} \right.$

$$y = h(x)$$

f_x

$$f_y = \sum_x P_x(x)$$

Theorem 3.1

(a) Let X be a discrete random variable and $h: \mathbb{R} \rightarrow \mathbb{R}$ is a function such that $E(h(X))$ exists then $E(h(X)) = \sum_x h(x) P_x(x)$

(b) Let X and Y be jointly discrete and $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ s.t $E(h(X, Y)) = \sum_x \sum_y h(x, y) P_{X,Y}(x, y)$

Ex) A r.v X has a pmf:

x	P _{x(y)}
1	0.5
2	0.5

a) find $E(X)$

b) Find $E\left[\frac{1}{X}\right]$

c) $y = \frac{1}{x} = h(x)$

a) $E(X) = (1 \times 0.5 + 2 \times 0.5) = 1.5 = \frac{3}{2}$

b) $E\left[\frac{1}{X}\right] = \sum_x \frac{1}{x} P_x(x) = \frac{1}{1} \cdot 0.5 + \frac{1}{2} \cdot 0.5 = 0.75$

Note:
 $\frac{1}{E(X)} = \frac{1}{\frac{3}{2}} = 0.67 \neq E\left[\frac{1}{X}\right]$

$$b) E\left[\frac{1}{x}\right] = \sum_x \frac{1}{x} P_X(x) = \frac{1}{1} \cdot 0.5 + \frac{1}{2} \cdot 0.5 = 0.75$$

Ex Random variables x and y are jointly discrete with joint pmf

$x \setminus y$	0	1
0	0.2	0.1
1	0.4	0.3
0.6 0.4		

a) Find $E(XY)$ { Ex: Find $E\left[\frac{X}{Y+1}\right]$

b) Find $E(x)$

c) Find $E(y)$

Answ a) $E(XY) = 0 \cdot 0 \cdot 0.2 + 0 \cdot 1 \cdot 0.1 + 1 \cdot 0 \cdot 0.4 + 1 \cdot 1 \cdot 0.3 = 0.3$

b) $E(x) = \sum x P_X(x) = 0 \cdot 0.3 + 1 \cdot 0.7 = 0.7$

c) $E(y) = \sum y P_Y(y) = 0 \cdot 0.6 + 0.4 = 0.4$

Note: $E(Y) \times E(X) \neq E(XY)$

St Petersburg Paradox

You toss a fair coin until you get a first head, if x is the # of tails before until the first head. You win $\$2^x$. What is the expected value of the award?

Answ $x \sim Geom(\theta = \frac{1}{2}) \Rightarrow P_X(x) = (1-\theta)^x \theta, x=0,1,2 \dots$ $P_X(x) = \left(\frac{1}{2}\right)^{x+1}, x=0,1,2 \dots$

$$E[2^x] = \sum_x 2^x P_X(x) = \sum_{x=0}^{\infty} 2^x \left(\frac{1}{2}\right)^{x+1}$$

$$= \sum_{x=0}^{\infty} \frac{1}{2} = \infty$$

b) $P(L \geq \$256) = P(2^x \geq 2^8) = P(x \geq 8) = P_X(8) + P_X(9) + P_X(10) + \dots$

$$= \left(\frac{1}{2}\right)^9 + \left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right)^{11} + \dots, \text{ recall } \frac{a}{1-r}$$

$$= \frac{\left(\frac{1}{2}\right)^9}{1 - \frac{1}{2}} = \left(\frac{1}{2}\right)^8 = 0.0039$$

Theorem 3.1.2 Linearity of Expectation

If x, y are jointly discrete with joint pmf P_{xy} then
 $E[ax + by] = E(x) + bE(y)$

$$\begin{aligned} \text{Pf } E[ax + by] &= \sum_x \sum_y (ax + by) P_{xy}(x, y) \\ &= \sum_x \sum_y (ax + aP_{xy}(x, y) + by P_{xy}(x, y)) \\ &= \sum_x ax \sum_y P_{xy}(x, y) + \sum_y by \sum_x P_{xy}(x, y) \\ &= a \sum_x x P_x(x) + b \sum_y y P_y(y) \\ &\stackrel{\text{By definition}}{=} aE(x) + bE(y) \end{aligned}$$

This also implies,

- ① $E(x+y) = E(x) + E(y)$, but product **No!**
- ② $E[ax] = aE(x)$



Last hour: (3.1 Expectation)

- Definition $E(x) = \sum_x x p_x(x)$

- $x \sim \text{Degenerate}(c) \Rightarrow E(x) = c$ $\bullet x \sim P_0(\lambda) \Rightarrow E_x = \lambda$

- $x \sim \text{Ber}(\theta) \Rightarrow E_x = \theta$ $\bullet x \sim G_p(\theta) \Rightarrow E_x = \frac{1-\theta}{\theta}$

- $x \sim \text{Bin}(n, \theta) \Rightarrow E_x = n\theta$

- $E(x)$ can be ∞ ($\pm \infty$)

- $E(x)$ can be non-existent (i.e. $\infty - \infty$)

- $E(h(x)) = \sum_x h(x) p_x(x)$... i.e. we do not need the pmf of $h(x)$ to calculate $E(h(x))$

- E is a linear operator

i.e. $E(ax+by) = aE(x) + bE(y)$ $\forall a, b \in \mathbb{R}$

$$\Rightarrow E(ax) = aE(x) \dots (\text{taking } b=0)$$

$$\Rightarrow E(x+b) = E(x) + b \quad (\text{taking } a=1 \text{ and } y=1)$$

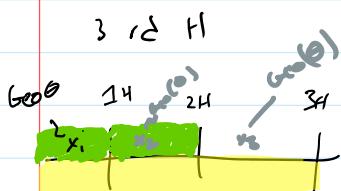
$$\Rightarrow E(x+y) = E(x) + E(y)$$

$$\Rightarrow E\left[\sum_{i=1}^k E(x_i)\right] = \sum_{i=1}^k E(x_i)$$

Notes

$$E(xy) \neq E(x)E(y)$$

Example Let $y \sim \text{Negbin}(r, \theta)$. Find $E(y)$



$x_1 \neq x_2 \neq x_3$ however distributions are identical.

We can say negative binomial is the sum of r geometric.

∴ when y is negative binomial : $\text{iid} \Rightarrow \text{independent and identically distributed}$

$$Y \sim \text{Neg bin}(r, \theta) \Rightarrow y = \sum_{i=1}^r x_i, \text{ where } x_i \stackrel{\text{iid}}{\sim} \text{Geo}(\theta)$$

$$E(y) = E\left[\sum_{i=1}^r x_i\right] = \sum_{i=1}^r E(x_i) = \sum_{i=1}^r \frac{1-\theta}{\theta} = r \left(\frac{1-\theta}{\theta}\right)$$

Recall

$$\sum_{i=1}^r k = nk$$

Theorem if x and y are independent then

Thrm if x and y are independent then
 $E[g(x) \cdot h(y)] = (E(g(x))) (E(h(y)))$

Proof (Discrete case)

$$\begin{aligned}
 E[g(x) \cdot h(y)] &= \sum_x \sum_y g(x) h(y) p_{x,y}(x,y) \\
 &= \sum_x \sum_y g(x) h(y) p_x(x) \cdot p_y(y) \\
 &= \sum_x g(x) p_x(x) \sum_y h(y) p_y(y) \\
 &= \sum_y h(y) p_y(y) \sum_x g(x) p_x(x) \\
 &= E(g(x)) \cdot E(h(y))
 \end{aligned}
 \quad \left. \begin{array}{l} \text{Recall} \\ E(g(x,y)) = \sum_x \sum_y g(x,y) p_{x,y}(x,y) \\ x \perp y \Rightarrow p_{x,y}(x,y) = p_x(x) p_y(y) \end{array} \right\}$$

Note if $x \perp y \Rightarrow E(x \cdot y) = E(x) E(y)$ ① only if independent

$$E(x+y) = E(x) + E(y) \quad \text{always true} \checkmark$$

Thrm 3.1.4 Monotonicity of expectations

$$\text{if } x \leq y \Rightarrow E(x) \leq E(y)$$

Proof

Define: $z = y - x$ then

$$z(s) = y(s) - x(s) \geq 0 \quad \forall s \in S$$

$$x \leq y \Leftrightarrow x(s) \leq y(s) \quad \forall s \in S$$

Define: $Z = Y - X$ then

$$Z(s) = Y(s) - X(s) \geq 0 \quad \forall s \in S$$
$$E(Y-X) = E(Z) = \sum_{z \geq 0} z \underbrace{P(Z=z)}_{P(Z=z) \geq 0}$$
$$\left| \begin{array}{l} P_Z(z) = P(Z=z) \\ \end{array} \right.$$

$$\therefore \sum_{z \geq 0} z P(Z=z) \geq 0$$

By linearity of E , $E(Y) - E(X) \geq 0$

$$\Rightarrow E(X) \leq E(Y)$$

3.2 Abs continuous Random variables

Definition (Expectation)

If X is an absolutely continuous random variable then its expectations is given by:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$
$$\left| \begin{array}{l} \text{Discrete case} \\ E_x = \sum_x x p_x(x) \\ = \int x f_X(x) dx \end{array} \right.$$

Ex $X \sim \text{Uniform}[L, R]$, find $E(X)$?

Recall $f_X(x)$ of uniform

$$f_X(x) = \begin{cases} \frac{1}{R-L}, & L \leq x \leq R \leftarrow \text{support} \\ 0, & \text{o/w} \end{cases}$$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

$$= \int_L^R x \frac{1}{R-L} dx$$

$$\begin{aligned}
 &= \int_L^R x \frac{1}{R-L} dx \\
 &= \frac{1}{R-L} \int_L^R x dx \\
 &= \frac{1}{R-L} \left[\frac{x^2}{2} \right]_L^R \\
 &= \frac{1}{R-L} \left(\frac{R^2 - L^2}{2} \right) = \frac{L+R}{2}
 \end{aligned}$$

Ex2. $x \sim \exp(\lambda)$. Find $E(x)$?

Ans

$$\begin{aligned}
 E(x) &= \int_0^\infty x f_x(x) dx = \int_0^\infty x \lambda e^{-\lambda x} dx \\
 u &= x \quad du = dx \\
 v &= -e^{-\lambda x} \quad dv = \lambda e^{-\lambda x} \\
 &\left[-x e^{-\lambda x} \right]_0^\infty + \int_0^\infty \lambda e^{-\lambda x} dx \\
 &= 0 + \frac{1}{\lambda} = \frac{1}{\lambda}
 \end{aligned}
 \quad \left| \begin{array}{l} \text{Recall} \\ f_x(x) = \lambda e^{-\lambda x}, x \geq 0 \\ = 0, \text{ otherwise} \end{array} \right.$$

$$x \sim \exp(\lambda) \Rightarrow E(x) = \frac{1}{\lambda}$$

Example 3 $z \sim N(0, 1)$. Find $E(z)$

Ans:

$$\begin{aligned}
 E(z) &= \int_{-\infty}^{\infty} z f_z(z) dz \\
 &= \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
 &= \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^0 z e^{-\frac{z^2}{2}} dz + \int_0^{\infty} z e^{-\frac{z^2}{2}} dz \right] \\
 &= \frac{1}{\sqrt{2\pi}} \left[\left[-e^{-\frac{z^2}{2}} \right]_{-\infty}^0 + \left[-e^{-\frac{z^2}{2}} \right]_0^{\infty} \right]
 \end{aligned}
 \quad \left| \begin{array}{l} \text{Recall} \\ f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, -\infty < z < \infty \end{array} \right.$$

$$\begin{aligned}
 &= \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^{-1} e^{-\frac{x^2}{2}} dx + \int_{-1}^{0} e^{-\frac{x^2}{2}} dx \right] \\
 &= \frac{1}{\sqrt{2\pi}} [(-1 - 0) + (0 - (-1))] \\
 &= 0
 \end{aligned}$$

$$z \sim N(\mu=0, \sigma^2=1) \Rightarrow E(z) = 0$$

Theorem 3.2.1

(a) if x is an absolute continuous with pdf f_x and $g: \mathbb{R} \rightarrow \mathbb{R}$ is a function, s.t $E(g(x))$ exists, then

$$E(g(x)) = \int g(x) \cdot f_x(x) dx \quad \left| \begin{array}{l} y = g(x) \\ E_y = \int y f_g(y) dy \end{array} \right.$$

(b) If x and y are jointly abs continuous with joint pdf $f_{x,y}$ and $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ s.t $E(h(x,y))$ exists then

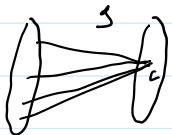
$$E[h(x,y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) f_{x,y}(x,y) dx dy \quad \left| \begin{array}{l} y = h(x,y) \\ E_y = \int y f_g(y) dy \end{array} \right.$$

Theorem 3.2.2 (Linearity of expectation)

If x & y are jointly absolute continuous then $E[\alpha x + by] = aE(x) + bE(y) \quad \forall a, b \in \mathbb{R}$

$$\begin{aligned}
 \Rightarrow E[\alpha x] &= \alpha E(x) \\
 \Rightarrow E[x+y] &= E(x) + E(y)
 \end{aligned}$$

Ex x is a absolutely continuous random variable with pdf f_x



$$g(x) = c$$

Find $E(g(x))$?

Soln

$$\begin{aligned} E[g(x)] &= \int_{-\infty}^{\infty} g(x) f_x(x) dx \\ &= \int_{-\infty}^{\infty} c f_x(x) dx \\ &= c \int_{-\infty}^{\infty} f_x(x) dx \\ &\quad \text{1 bc cause its a density function} \\ &\quad \text{and by def } \int_{-\infty}^{\infty} f_x(x) dx = 1 \\ &= c \end{aligned}$$

$$\therefore E(c) = c$$

Ex $x \sim N(\mu, \sigma^2)$. Find $E(x)$

$$\text{Soln } x \sim N(\mu, \sigma^2) \Rightarrow z = \frac{x-\mu}{\sigma} \sim N(0, 1)$$

$$0 = E(z) = E\left[\frac{x-\mu}{\sigma}\right]$$

$$- \quad - \quad - \frac{\mu}{\sigma}$$

$$= \frac{1}{\sigma} E(x) - \frac{\mu}{\sigma} \Rightarrow E(x) = \mu$$

Ex Find pdf

of a function

of a given

r.v

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

$$z = \frac{x-\mu}{\sigma} = h(x)$$

find the distribution
of z .

Hints: Find the pdf of
 z . Identify the
distribution

$$\left[\text{Ans: } \frac{x-\mu}{\sigma} \sim N(0, 1) \right]$$

ANSWER

$$\left[\text{Ans: } \frac{x-\mu}{\sigma} \sim N(0,1) \right]$$

$$\begin{cases} E_x X \sim f_x(x) = \frac{1}{x^2}, & x > 1 \\ = 0, & \text{otherwise} \end{cases}$$

show it's a pdf
 $\frac{\oplus}{\ominus}$ and $\int_{-\infty}^{\infty} f_x(x) dx = 1$

Find $E(x)$?

$$\begin{aligned} \text{So } E(x) &= \int_{-\infty}^{\infty} x f_x(x) dx \\ &= \int_1^{\infty} x \cdot \frac{1}{x^2} dx = \int_1^{\infty} \frac{1}{x} dx = [\ln|x|]_1^{\infty} \\ &= \infty - 0 = \infty \end{aligned}$$

$$E_x X \sim f_x(x) = \begin{cases} \frac{1}{2x^2}, & x > 1 \\ \frac{1}{2x^2}, & x < -1 \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} \text{Find } E(x) &= \int_{-\infty}^{-1} x f_x(x) dx + \int_1^{\infty} x f_x(x) dx \\ &= \frac{1}{2} \int_{-\infty}^{-1} x \cdot \frac{1}{2x^2} dx + \frac{1}{2} \int_1^{\infty} x \cdot \frac{1}{2x^2} dx \\ &= \frac{1}{2} \left[\ln|x| \right]_{-\infty}^{-1} + \frac{1}{2} \left[\ln|x| \right]_1^{\infty} \\ &= \frac{1}{2} [0 - \ln(-\infty)] + \frac{1}{2} [\ln(\infty) - 0] \\ &= \frac{1}{2} [0 - \infty] + \frac{1}{2} [\infty - 0] \\ &= -\infty + \infty = \infty - \infty \Rightarrow \text{DNE} \end{aligned}$$

Theorem 3.2.3

If $x \perp y$, then $E[g(x)h(y)] = E(g(x))E(h(y))$

In particular this $\Rightarrow E(xy) = E(x) \cdot E(y)$

Thm 3.2-4

$$x \leq y \Rightarrow E(x) \leq E(y)$$

3.3 Variance, Covariance, Correlation

Definition (Variance)

The variance of a random variable x is defined by
 $E[(x - E(x))^2]$ Notation. $E(x) = \mu_x$ with $\text{Var}(x) = E[(x - \mu_x)^2]$

Note: $\text{Var}(y) \geq 0$

Result: $\text{Var}(x) = E[x^2] - [E(x)]^2$ or $(E(x^2) - \mu_x^2)$

PF $\text{Var}(x) = E[(x - \mu_x)^2]$

$$= E[x^2 - 2\mu_x x + \mu_x^2]$$

$$= E(x^2) - 2\mu_x E(x) + \mu_x^2$$

\uparrow
 x

$$= E(x^2) - 2\mu_x^2 + \mu_x^2$$

$$= E(x^2) - \mu_x^2 = E(x^2) - (E(x))^2$$

Note $(E(x))^2 \leq E(x^2)$ because $\text{var}(x) \geq 0$

Last hour: Section 3.3 variance

Definition: $\text{Var}(x) = E[(x - E(x))^2] \leftarrow \text{harder}$

$$\begin{cases} E(x) \\ E(x^2) \end{cases}$$

Results $\text{Var}(x) = E(x^2) - (E(x))^2 \leftarrow \text{easier}$

Notes: $\text{Var}(x) \geq 0 \quad \forall \text{ r.v. } x$
 $(E(x))^2 \leq E(x^2), \quad \forall \text{ r.v. } x$

Before var

- $x \perp y \Rightarrow E[g(x)h(y)] = E(g(x)) \cdot E(h(y))$

$$(x \perp x \Rightarrow E(xy) = Ex \cdot Ey)$$

- $x \leq y \Rightarrow E(x) \leq E(y) \quad \text{monotonicity of expectation}$
- $E(ax+by) = aE(x) + bE(y) \quad \text{linearity}$

Definition: Standard deviation of a r.v. x

$$\text{SD}(x) = \sqrt{\text{Var}(x)}, \quad (\text{Denoted by } \sigma_x)$$

Results

① $\text{Var}(c) = 0$, since a constant doesn't vary. It is fixed.

pf: $\text{Var}(c) = E[(c - E(c))^2]$

$$\begin{aligned} &= E[(c - c)^2] \\ &= E[0] = 0 \end{aligned}$$

② $\text{Var}(ax + bx) = b^2 \text{Var}(x)$

$$\begin{aligned}
 \text{PF } \text{Var}(a+bx) &= E[(a+bx) - E(a+bx)^2] \\
 &= E[(a+bx - a - bE(x))^2] \\
 &= E[b(x-E(x))^2] \\
 &= b^2 E[(x-E(x))^2] \\
 &= b \text{Var}(x)
 \end{aligned}$$

This result implies:

- $\text{Var}(a+x) = \text{Var}(x)$
- $\text{Var}(bx) = b^2 \text{Var}(x)$ ← remainder of variance when you take a constant out, to square it.
- $\text{SD}(a+bx) = \sqrt{\text{Var}(a+bx)}$
 $= \sqrt{b^2 \text{Var}(x)}$
 $= |b| \text{SD}(x)$
 $= |b| \sqrt{\text{Var}(x)}$

Ex $x \sim \text{Ber}(\theta)$, Find $\text{Var}(x)$

Ans

$$\begin{aligned}
 \text{Var}(x) &= E(x^2) - (E(x))^2 \\
 &= \theta - (\theta)^2 \\
 &= \theta(1-\theta)
 \end{aligned}$$

$$\therefore x \sim \text{Ber}(\theta)$$

$$\Rightarrow \text{Var}(x) = \theta(1-\theta)$$

$$E(x) = \theta$$

Recall

$$E(x) = \theta \text{ for binomial}$$

$$P_x(x)$$

x	0	1
$P_x(x)$	θ	$1-\theta$
1	0	0

$$E(x^2)$$

$$\begin{aligned}
 &= \sum_x x^2 P_x(x) \\
 &= 0^2 \cdot (1-\theta) + 1^2 \cdot \theta \\
 &= \theta
 \end{aligned}$$

Ex $x \sim \text{Bin}(n, \theta)$. Find variance

$$\text{Ans } \text{Var}(x) = E(x^2) - (E(x))^2$$

$$\begin{aligned} E[x(x-1)] &= \sum_{x=0}^n x(x-1) \binom{n}{x} \theta^x (1-\theta)^{n-x} \\ &\quad \text{Some can change this to two} \\ &= \sum_{x=2}^n x(x-1) \frac{n!}{x!(n-x)!} \theta^x (1-\theta)^{n-x} \\ &= \sum_{x=2}^n (x-1) \frac{n!}{(x-2)!(n-x)!} \theta^x (1-\theta)^{n-x} \\ &= n(n-1)\theta^2 \sum_{x=2}^n \frac{(n-2)!}{(n-2)![x(n-2)-x]!} \theta^{n-2} (1-\theta)^{x(n-2)-x} \\ &= n(n-1)\theta^2 \sum_{y=0}^{n-2} \frac{(n-2)!}{y!(n-2-y)!} \theta^y (1-\theta)^{(n-2)-y} \\ &\quad \text{binomial} \\ &= (\theta + 1-\theta)^{n-2} \cdot n(n-1)\theta^2 \\ &= n^2\theta^2 - n\theta^2 \end{aligned}$$

$$E[x(x-1)] = n^2\theta^2 - n\theta^2$$

$$E(x^2) - E(x) = n^2\theta^2 - n\theta^2$$

$$E(x^2) = n^2\theta^2 - n\theta^2 + n\theta$$

$$\begin{aligned} \text{Var}(x) &= E(x^2) - (E(x))^2 \\ &= n^2\theta^2 - n\theta^2 + n\theta - (n\theta)^2 \\ &= n\theta(1-\theta) \end{aligned}$$

$$x \sim B(n, \theta)$$

$$\Rightarrow \text{Var}(x) = n\theta(1-\theta)$$

$$\Rightarrow E(x) = n\theta$$

Recall

if $x \sim \text{bin}$ then

$$E(x) = n\theta$$

$$\sum g(x) P_x(x)$$

$$P_x = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

$$\text{Sub } y = x-2$$

$$(a+b)^n = \sum \binom{n}{r} a^r b^{n-r}$$

$$a = \theta$$

$$b = 1-\theta$$

recall $E(x) = n\theta$ for
binomial

Ex $x \sim P_o(\lambda)$, Find $\text{Var}(x)$

Ex $X \sim P_0(\lambda)$, Find $\text{Var}(X)$

Ans $\text{Var}(X) = E(X^2) - (E(X))^2$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 p_x(x) \leftarrow \begin{array}{l} \text{too many} \\ \text{use Factorial} \\ \text{shortcuts} \end{array}$$

$$\begin{aligned} E[X(X-1)] &= \sum_{x=2}^{\infty} x(x-1) e^{-\lambda} \frac{\lambda^x}{x!} \\ &= e^{-\lambda} \lambda^2 \underbrace{\left[\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} \dots \right]}_{e^\lambda} \end{aligned}$$

$$\begin{aligned} &= e^{-\lambda} \lambda^2 \\ &= \lambda^2 \end{aligned}$$

Recall

If $X \sim P_0$

$E(X) = \lambda$

$p_x(x) = e^{-\lambda} \frac{\lambda^x}{x!}$

change summation limits

use factorial shortcut

convert all to $(\lambda - x)$ form

sub $y = (x-2)$
(not in this case though)

$$E(X(X-1)) = \lambda^2$$

$$E(X^2) - E(X) = \lambda^2$$

$$E(X^2) - \lambda = \lambda^2$$

$$\begin{aligned} E(X^2) &= \lambda^2 + \lambda \\ &= \lambda(\lambda + 1) \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= E(X^2) - (E(X))^2 \\ &= \lambda^2 + \lambda - \lambda^2 \\ &= \lambda \end{aligned}$$

\therefore if $X \sim P_0(\lambda) = \text{Var}(X) = \lambda$

$$E(X) = \lambda$$

if mean and variance =
use poisson.

Last hour: Sec 3.3 Variance

- * Definition: $\text{Var}(x) = E[(x - \mu_x)^2]$ where $\mu_x = E(x)$ ← Difficult to calculate variance with the definition, use the result.
- Result: $\text{Var}(x) = E(x^2) - (E(x))^2$
- Standard deviation, $\text{SD}(x) = \sqrt{\text{Var}(x)}$
- Properties
 - $\text{Var}(x) > 0$
 - $\text{Var}(ax + b) = a^2 \text{Var}(x)$ ← remember to square the a^2
 - $\text{Var}(c) = 0 \quad \forall c \in \mathbb{R}$

Examples

$$\begin{aligned} x \sim \text{Ber}(\theta) \Rightarrow \text{Var}(x) = \theta(1-\theta), \quad E(x) = \theta \\ x \sim \text{Bin}(n, \theta) \Rightarrow \text{Var}(x) = n\theta(1-\theta), \quad E(x) = n\theta \\ x \sim P_0(\lambda) \Rightarrow \text{Var}(x) = \lambda, \quad E(x) = \lambda \end{aligned}$$

Today

Ex/ $x \sim \text{Geo}(\theta)$. Find $\text{Var}(x)$?

$$\text{Ans} \quad V(x) = E(x^2) - (E(x))^2$$

$$\begin{aligned} E(x(x-1)) &= \sum_{x=0}^{\infty} x(x-1)(1-\theta)^x \theta \\ &= \sum_{x=2}^{\infty} x(x-1)(1-\theta)^x \theta \\ &= \theta \sum_{x=2}^{\infty} (x)(x-1)(q)^x, \quad \text{let } q = 1-\theta \\ &= \theta q^2 \sum_{x=2}^{\infty} x(x-1)q^{x-2} \\ &= \theta q^2 \sum_{x=2}^{\infty} x(x-1)q^{x-2} \\ &= \theta q^2 \sum_{x=2}^{\infty} \frac{d^2}{dq^2} q^x \\ &= \theta q^2 \frac{d^2}{dq^2} \sum_{x=2}^{\infty} q^x \leftarrow \text{geometric series} \\ &= \theta q^2 \frac{d^2}{dq^2} \left(\frac{q^2 - 1}{1-q} \right) \end{aligned}$$

$$= \theta q^2 \frac{d^2}{dq^2} \left[-1 - q + \frac{1}{1-q} \right]$$

Hint

$E x^2$ is difficult
use $E(x(x-1))$

helps when (!) \leftarrow factorials
are involved.

Recall

if $X \sim \text{Geo}(\theta)$ then

$$E(x) = \frac{1-\theta}{\theta} \quad p_X(x) = (1-\theta)^{x-1} \theta, \quad x=0, 1, 2, \dots$$

Recall

$$\sum_{x=2}^{\infty} x(x-1)q^{x-2} \leftarrow \text{power series}$$

Interchange summation and derivatives

$$= \theta q^2 \left[\frac{2}{(1-q)^3} \right]$$

recall, $q = 1-\theta$

$$= \theta q^2 \left[\frac{2}{\theta^2} \right] = \frac{2q^2}{\theta^2}$$

$$= \frac{2(1-\theta)^2}{\theta^2}$$

$$E(x^2) = \frac{2(1-\theta)^2}{\theta^2} + \frac{1-\theta}{\theta}$$

$$V(x) = E(x^2) - (E(x))^2$$

$$= \frac{2(1-\theta)^2}{\theta^2} + \frac{1-\theta}{\theta} - \frac{(1-\theta)^2}{\theta^2}$$

$$= \frac{(1-\theta)^2}{\theta} + \frac{1-\theta}{\theta} = \frac{1-\theta}{\theta} ((1-\theta) + 1)$$

$$= \frac{1-\theta}{\theta^2}$$

\therefore if $x \sim \text{Geo}(\theta)$ then

$$\Rightarrow V(x) = \frac{1-\theta}{\theta^2} \quad E(x) = \frac{1-\theta}{\theta}$$

Ex $x \sim \text{Exp}(\lambda)$ show that $V(x) = \frac{1}{\lambda^2}$

Hints $V(x) = E(x^2) - (E(x))^2$

$$E(x^2) = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}, \text{ show this}$$

$$u = \lambda x \quad du = x dx$$

$$= \int_0^\infty \frac{u}{\lambda} \lambda^{-u} du$$

$$= \int_0^\infty \frac{u}{\lambda} \frac{1}{\lambda^u} du \quad (\text{homework exercise})$$

$$= \frac{1}{\lambda} \int_0^\infty \frac{u}{\lambda^u} du$$

Recall

$$E(x) = \frac{1}{\lambda}$$

$$f_x(x) = \lambda e^{-\lambda x} \quad x > 0$$

Ex / $z \sim N(0, 1)$. Find $V(z)$

$$V(z) = E(z^2)$$

$$= \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$= \int_{-\infty}^{\infty} z \cdot z \cdot \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}_{dv} dz$$

$$\left. \begin{array}{l} E(z) = 0 \\ V(z) = E(z^2) - (E(z))^2 \\ \Rightarrow V(z) = E(z^2) - 0 = E(z^2) \end{array} \right\}$$

$$\text{Let } u = z, dv = \frac{1}{\sqrt{2\pi}} z \cdot e^{-\frac{1}{2}z^2} dz$$

$$du = 1, v = \frac{1}{\sqrt{2\pi}} \left(-e^{-\frac{1}{2}z^2} \right)$$

$$= \frac{1}{\sqrt{2\pi}} \left[-z^2 e^{-\frac{1}{2}z^2} \right]_{-\infty}^{\infty} + \int_0^{\infty} e^{-\frac{1}{2}z^2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = 1 \quad \leftarrow \text{since this is a density function}$$

the integral is $\int_0^{\infty} 1$

$$Z \sim N(\mu, \sigma^2) \Rightarrow V(z) = 1, E(z) = \mu$$

Recall $N(0, 1)$ is standard normal
and $N(\mu, \sigma^2)$ is normal.

$$\text{Ex } x \sim N(\mu, \sigma^2) \Rightarrow z = \frac{x - \mu}{\sigma} \sim N(0, 1)$$

$$\begin{aligned} z &= \frac{1}{\sigma} x - \frac{\mu}{\sigma} \\ 1 &= V(z) = \left(\frac{1}{\sigma^2}\right) V(x) \\ \Rightarrow V(x) &= \sigma^2 \end{aligned}$$

resembles $ax + b$
Recall
 $V(ax + b) = a^2 V(x)$
 $x \sim N(\mu, \sigma^2)$
 $E(x) = \mu, V(x) = \sigma^2$

Ex $x \sim \text{Gamma}(\alpha, \lambda)$. Find $V(x)$?

Ex $x \sim \text{Gamma}(\alpha, \lambda)$. Find $\text{Var}(x)$?

$$\text{Ans } V(x) = E(x^2) - (E(x))^2$$

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

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

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

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}}$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\alpha \Gamma(\alpha)}{\lambda^{\alpha+1}}$$

$$= \frac{\alpha}{\lambda}$$

$$\therefore E(x) = \frac{\alpha(\alpha+1)}{\lambda^2}$$

$$\text{Var}(x) = E(x^2) - (E(x))^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha^2}{\lambda^2}$$

Ex $x \sim \text{Beta}(\alpha, b)$. Show that

$$(1). E(x) = \frac{\alpha}{\alpha+b}$$

$$(2). \text{Var}(x) = \frac{\alpha b}{(a+b)^2(a+b+1)}$$

Recall

Density

$$f_x(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, x \geq 0$$

because it's a density function

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

$$f(x,y) = \begin{cases} \frac{1}{2} & , x \geq 0, y \geq 0, x+y \leq 2 \\ 0 & \text{o/w} \end{cases}$$

$$P(x>1 | y=0.5)$$

$$P_{x|y}(x|y) = \frac{P_{x,y}(x,y)}{P_y(y)}$$

$$P_{x|y}(x>1 | y=0.5) = \frac{\frac{1}{2}}{\frac{2}{2}}$$

$$\int_0^{\frac{2-y}{2}} \frac{1}{2} dx = \frac{1}{2} \cdot \frac{2-y}{2}$$

$$\frac{2-y}{2} = \frac{1}{2} \times \frac{x}{2-y}$$

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

$$= \frac{1}{\frac{3}{2}}$$

$$= \frac{2}{3}$$

$$C \int_0^{\infty} e^{-\frac{y}{1000}} x = 1$$

$$(C \left[C^{-\frac{y}{1000}}(x) \right] = C^{-\frac{y}{1000}}(6))$$

7

= /





c \int_0^x $-$ \int_0^y

$-$ $\{$ C $=$

C $-$ $\{$

$f(x) =$ t

— /

Up
— — X
— (200)

6

O H



$$\left. \begin{array}{c} e^{-\frac{4}{1000}x} \\ e^{-\frac{4}{1000}t} \end{array} \right\} \rightarrow e^{-\frac{4}{1000}(x-t)}$$

$$\left. \begin{array}{c} e^{-\frac{4}{1000}x} \\ e^{-\frac{4}{1000}t} \end{array} \right\} \rightarrow e^{-\frac{4}{1000}(x-t)}$$

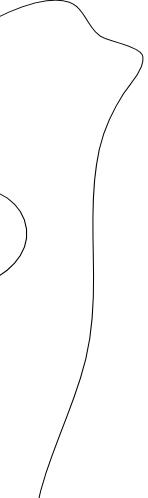
$$1 \times = 6.5$$

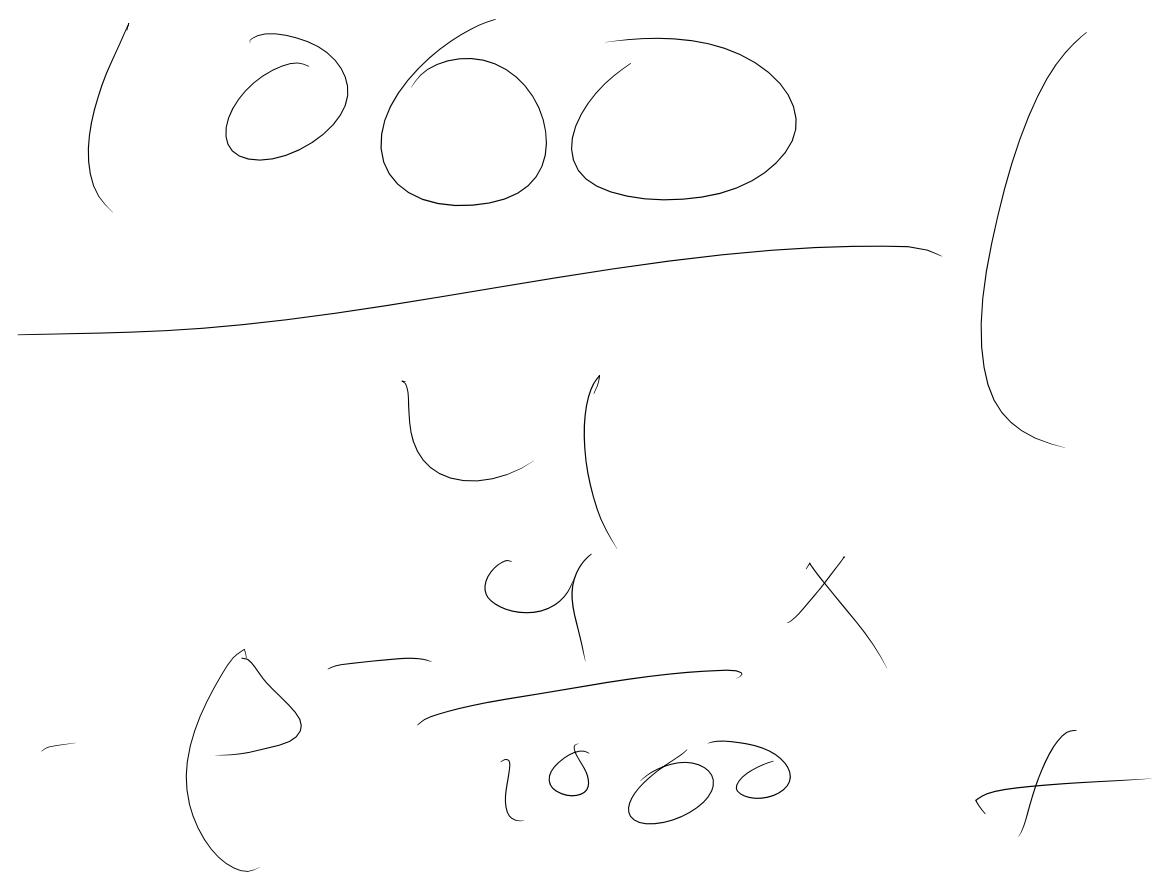
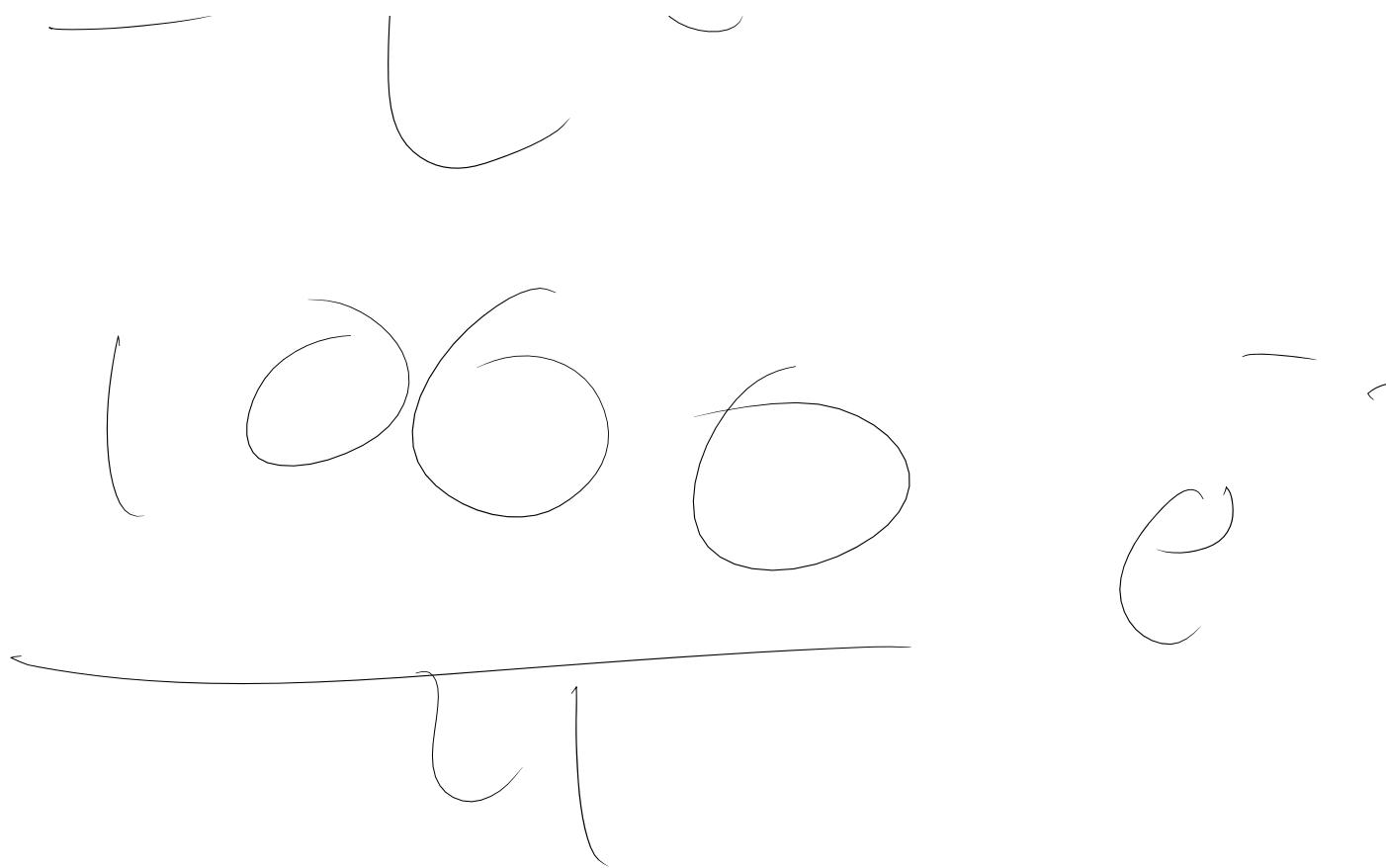
$$1 \times = 6.5$$

$$\begin{array}{r} 1006 \\ \hline - 4 \\ \hline \end{array} \quad \left. \begin{array}{l} \text{---} \\ \text{---} \end{array} \right\} \quad \begin{array}{r} 6 \\ \hline \end{array} \quad \begin{array}{r} 1005 \\ \hline \end{array}$$

$$\begin{array}{r} 1006 \\ \hline - 4 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ \hline \end{array} \quad \begin{array}{r} 1005 \\ \hline \end{array}$$

$$\begin{array}{r} 1005 \\ \hline - 4 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ \hline \end{array}$$





A sheet of white paper with horizontal blue lines. Handwritten in black ink:

- Top left: A horizontal line followed by two open parentheses, one above the other.
- Top right: A large open parenthesis followed by a vertical line.
- Middle left: The number 4.
- Middle center: A large open parenthesis followed by a plus sign.
- Middle left (continued): The number 100.
- Middle right: The number 1006.
- Bottom right: An open parenthesis followed by a vertical line.

$$= 4 \times 1000 + 4 = 4004$$

A collection of hand-drawn mathematical symbols and lines. On the left, there is a large vertical brace-like symbol. Next to it are two sets of wavy lines: one set of two parallel wavy lines and another set of three wavy lines. In the center, there is a large, irregular Z-shaped curve. Below the Z-shape is a long, thin, straight horizontal line. At the bottom right, there is a bracket symbol that spans the width of several small, rounded, looped curves.

(

→ O ↗

5 ↗

e^- ~~4~~ γ
 $(\partial\partial)$

e^- ~~cl~~ χ
 $(\partial\partial\partial)$

(82) - 4
- $(\partial\partial\delta)$
e ∂

1000

(

)

—

—

500

—

—

499

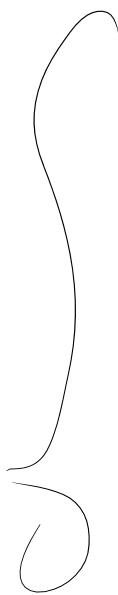
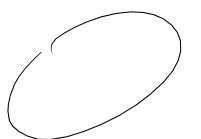
—

500

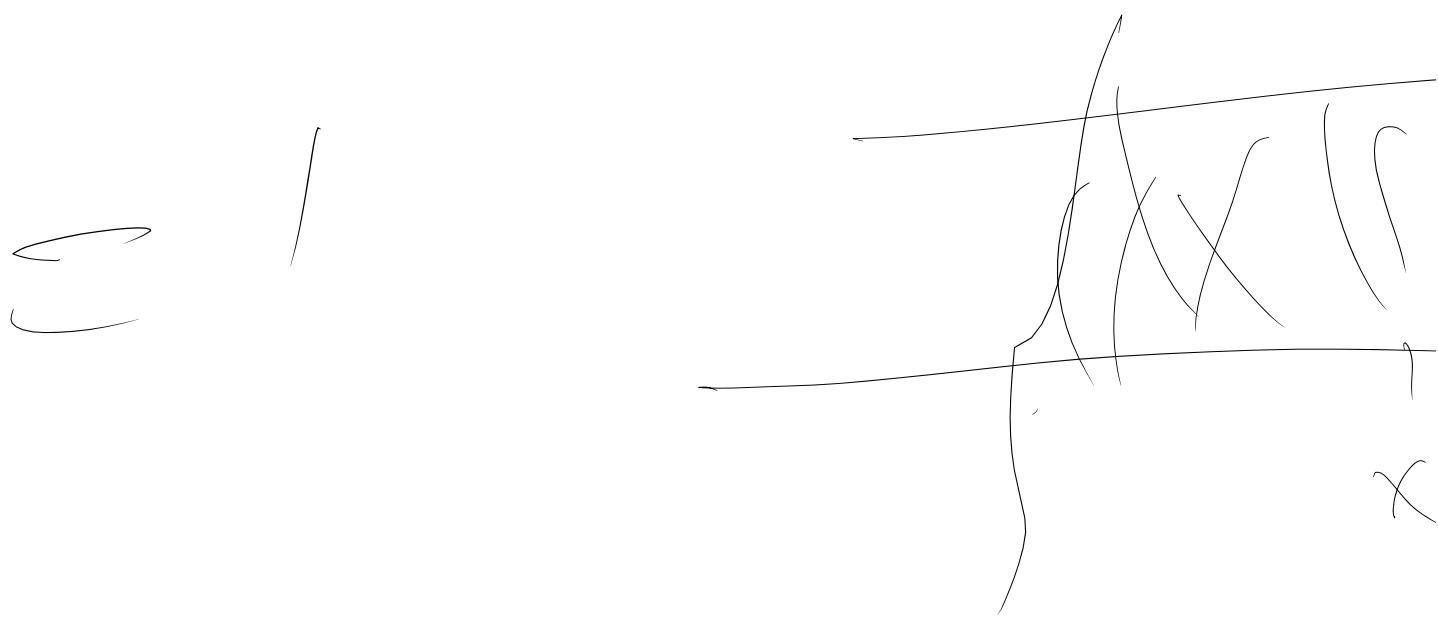
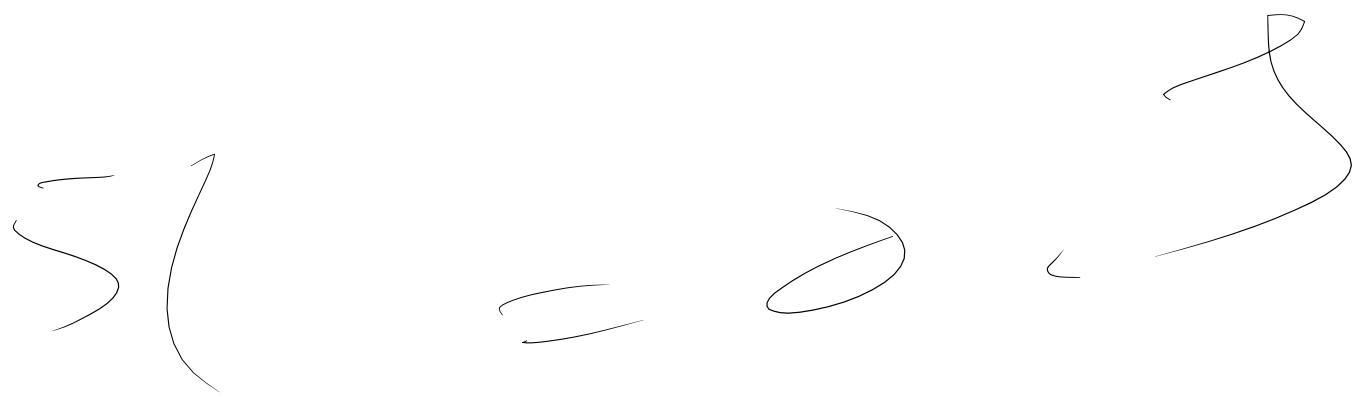
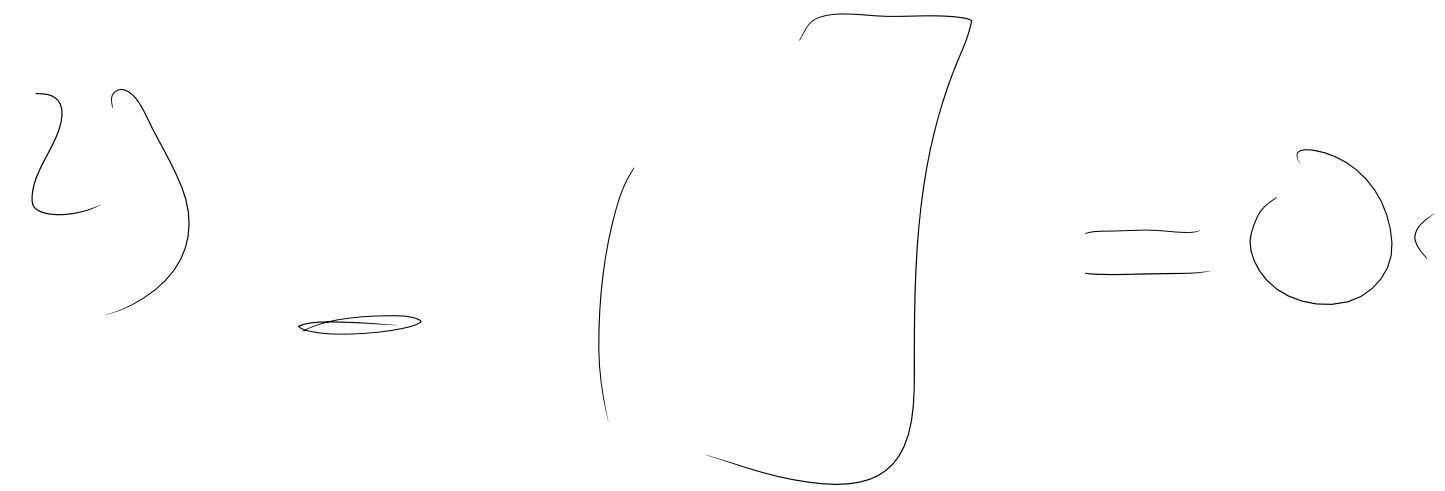
X

J X

$e^{-\frac{q}{100}} (c)$

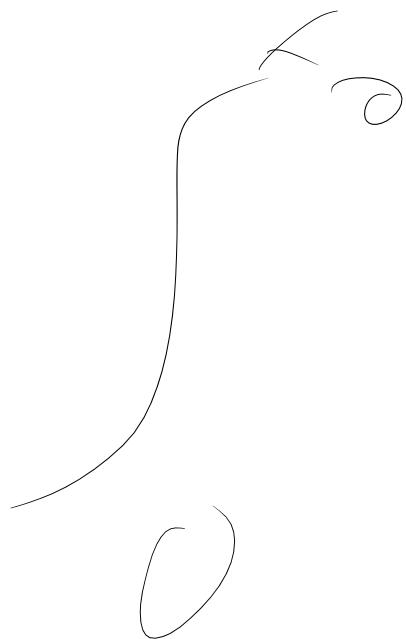
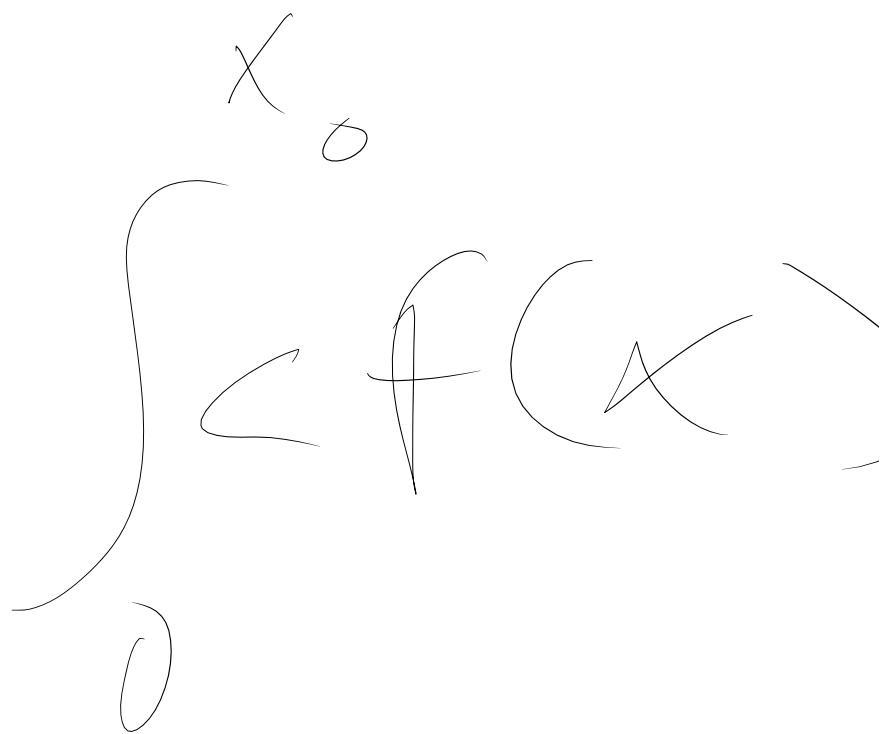


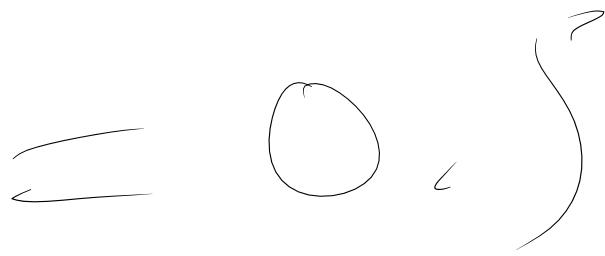
$f(\infty)$



~G











n_{white}

n_{black}

n_{red}

— — —

$$P(3w) = \frac{1}{15}$$

$$P(3w) = \underline{\quad}$$

$\underline{\quad}$

$$= \frac{n}{n}$$

$$\frac{1}{12} = \frac{n}{n-1}$$

$$\underline{\quad} = n(n-1)$$

$\sqrt{2}$

$\underline{12} =$

$$(2n)! = 2^n \cdot \underbrace{\dots}_{<}$$

$$\begin{aligned} & (2n)! \\ & \frac{n!}{(n-3)! 3!} \\ & \frac{(2n-1)!}{(2n-3)! 3!} \\ & \frac{n! (2n-3)!}{(2n)!} \\ & \frac{(-3)!(2n)!}{(-3)!(2n)!} \end{aligned}$$

$$-1)(n-2)(n-3) + \cancel{(2n-3)}$$

$$\cancel{-3} + (2n)(2n-1)(2n-2)(2n-3)$$



\sqrt{z}

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

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

$(z \in)$

\sqrt{z}

$$\begin{aligned}
 & \cancel{-3} + (2n)(n-1) - \frac{1}{6} = \\
 & (n-1)(n-2) \\
 & 2(2n-1)(2n-2) \\
 & \frac{(n-1)(n-2)}{2(2n-1)(2n-2)} \\
 & = n(n-1)(n-2) \quad 4n^2
 \end{aligned}$$

$$\int x f(x) dx =$$

$$\frac{(n-1)(n-2)}{(2n-1)(2n-2)}$$

$$] (2n-2) \geq (n-1)$$

$$-6n + 2 = 6$$

5 (x)



$$\int x^2 \left(\frac{1}{16} x^2 c \right) - 6.$$

$$P(x < y)$$

$$P(\angle 6)$$

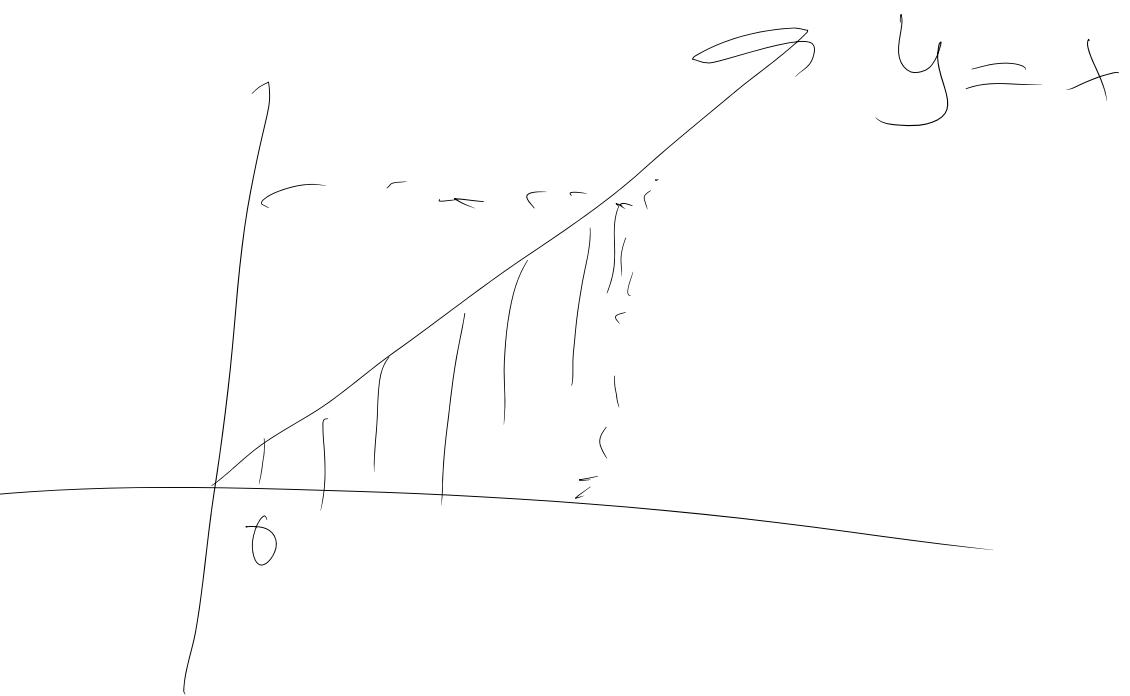
$$x$$

$$\int d \int d x \rightarrow x$$

$$x$$

S_x

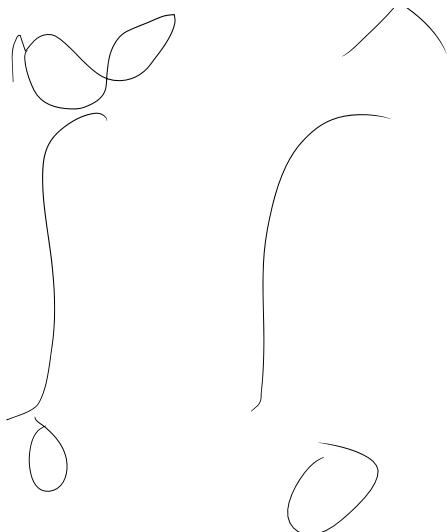
$$f(x) = \begin{cases} x & x > 0 \\ 0 & \text{otherwise} \end{cases}$$



- i ~ z







P C X G

P C O 2 L

P C - G , (

o - 2 < o - 3

o - 2 - o - 3 > 6

- o - | < 6

)

o - 3

6 o)



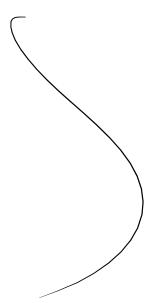
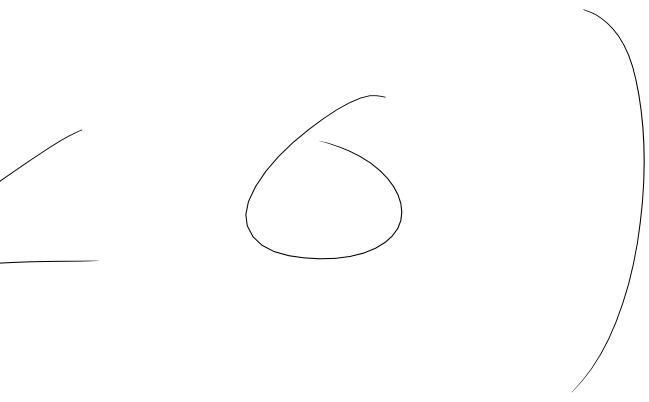


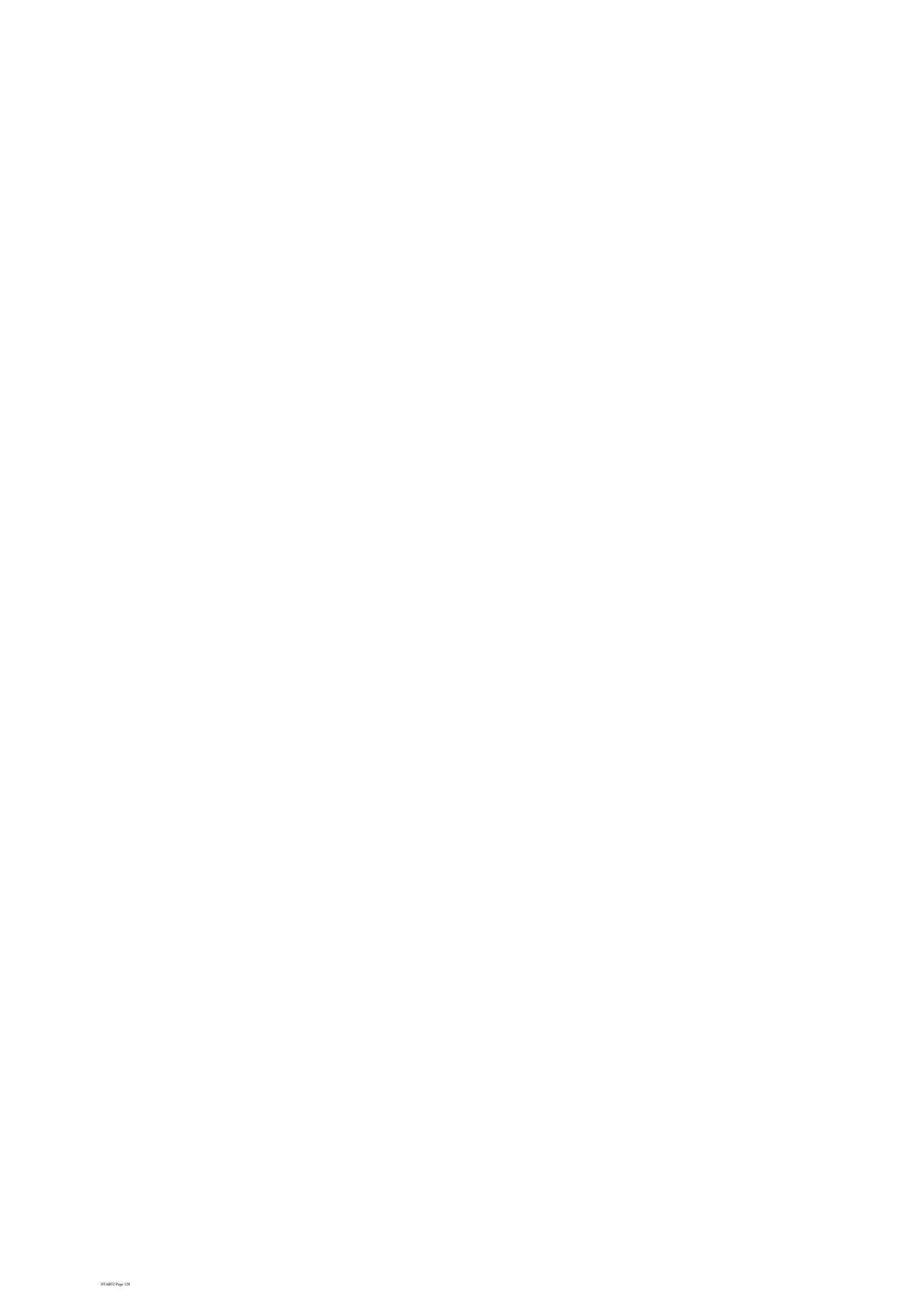
$P(x) = 0, \beta$

$f(x) > 0$

$E(x) = 0, \beta$

(





$$\int x (ax + b) \, dx$$

$$\int_0^1 ax^2 + bx \, dx$$

$$[ax^3]_0^1 = f(1) - f(0)$$

$$a \cancel{+} b' =$$

$$x^2) dx = \int$$

$$3 \quad dx = \int$$

$$\cancel{dx} \quad \cancel{da} \quad \left. \int \right|_U = \int$$

(



$$\frac{a}{3} + \frac{b}{4} =$$

$$a =$$

$$3$$

$$a = 3 -$$

$$a = 3(1)$$

=

1

4

35

$$\begin{array}{r} 4 \\ \underline{-} \\ 5 \end{array}$$

4

5





$$x(a x + 5)$$

(1)

$$\frac{ax^3}{3} + bx$$

$$\frac{a}{3} + \frac{5}{4}$$

$$x^2) = \textcircled{1}.$$

A series of hand-drawn mathematical symbols and numbers on a white background. The symbols include a large 'Y' with a horizontal line through it, a large curly brace, a vertical line with a small tick at the top, a large equals sign with a horizontal line through it, a large circle with a horizontal line through it, a small circle with a horizontal line through it, a large '6' with a horizontal line through it, and a small '6' with a horizontal line through it.





$$q = z$$

$$q = z$$

$$\alpha =$$

$$\left\{ \begin{array}{l} \\ ax + b \\ \end{array} \right.$$

$$\left[\frac{ax^2}{z} + b \right] =$$

$$\begin{array}{r} 2 - 4 = 6 \\ \hline 4 \end{array}$$

$$\begin{array}{r} 2 - 6 \\ \hline 6 \end{array}$$

$$x^2 = \int_a^b ax^2 + bx^3 dx$$

$$\left[\frac{ax^3}{3} + \frac{bx^4}{4} \right]_0^1$$

26

26



$$\begin{array}{r} x^2 \\ + k \\ \hline z \end{array}$$

$$3a +$$

$$a \leq$$

$$4(6 -$$

$$24 -$$

$$\frac{a}{3} + \frac{b}{4} = 0.6$$

$$4(a+3b) = 7.2$$

$$2b = 6$$

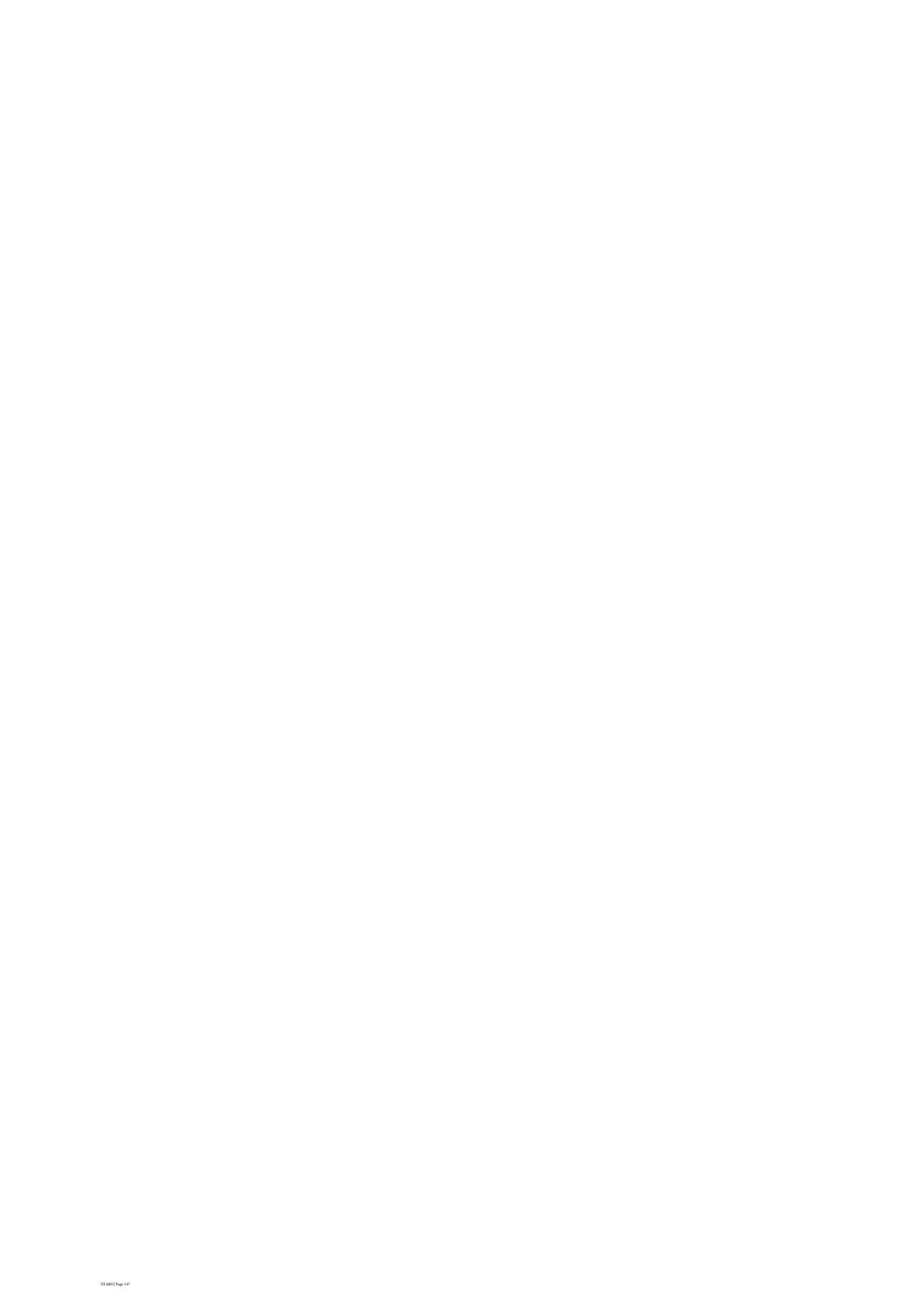
$$\underline{6 - 2b}$$

$$\underline{3}$$

$$\underline{2b}) + 3b = ? \quad |$$

$$\underline{8b} + 3b = ? \quad |$$

71 ✓





24 -

24

() }

$$8b + 3b = 21.6$$

$$-5b = 21.6$$

$$0.48 = b$$



Last hour (Covariance, Correlation)

Definition

$$\begin{aligned}\text{Cov}(x, y) &= E[(x - \mu_x)(y - \mu_y)] \\ &= E(xy) - E(x)E(y)\end{aligned}$$

easy way to calculate

Properties

- $\text{Cov}(x, y) = \text{Cov}(y, x)$
- $\text{Cov}(x, ax + by) = b\text{Cov}(x, y)$
- $\text{Cov}(x_1 + x_2, y) = \text{Cov}(x_1, y) + \text{Cov}(x_2, y)$
- $\text{Cov}(x, x) = V(x)$
- $V_{\text{ar}}(\sum_{i=1}^n x_i) = \sum_{i=1}^n V(x_i) + 2 \sum_{i=j} \sum \text{Cov}(x_i, x_j)$
- $\text{Corr}(x, y) = \frac{\text{Cov}(x, y)}{\sqrt{V(x)V(y)}}$
- if $y = a + bx$ then $\text{Corr}(x, y) = \frac{b}{|b|}$



Today 3.4 Generating functions

Probability generating function

Definition

$$r_x(t) = E[t^x] \quad \forall t \in \mathbb{R}$$

Moment generating function (m.g.f.)

Definition

$$M_x(t) = E[e^{tx}] \quad \forall t \in \mathbb{R}$$

← often used for discrete random variables.

Notes

$$\textcircled{1} \quad r_x(t) = E[t^x] \leftarrow \begin{array}{l} \text{definition} \\ \text{probability generating function} \end{array}$$

$$r_x(t) = E(t^x) = E(e^{t \ln t})$$

$$= E[e^{t \ln t}] = m_x(\ln t)$$

recall
 $t^x = e^{t \ln t}$

$$\textcircled{2} \quad m_x(t) = E[e^{tx}] = E[(e^t)^x] = r_x(e^t)$$

Ex $x \sim \text{Bin}(n, \theta)$. Find $r_x(t)$ and $m_x(t)$

$$r_x(t) = E[t^x] = \sum t^x p_x(x)$$

$$= \sum_{x=0}^n t^x \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

$$= \sum_{x=0}^n \binom{n}{x} (\theta t)^x (1-\theta)^{n-x}$$

$$= (1-\theta + \theta t)^n$$

recall
 $p_x(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$
 $x = 0, \dots, n$

binomial
 $(a+b)^n = \sum \binom{n}{x} a^x b^{n-x}$

\therefore for $\text{bin}(n, \theta)$

$$r_x(t) = (1-\theta + \theta t)^n \quad \forall t \in \mathbb{R}$$

$$m_x(t) = (1-\theta + \theta e^t)^n$$

Exercise, find it using first principle.

$$\text{Ex } m_x(t) = E(e^{tx}) = (1-\theta + \theta e^t)^n \quad \text{show this.}$$

$$\text{Ex } x \sim P_0(\lambda), \text{ show that } r_x(t) = e^{\lambda(t-1)}, \quad t \in \mathbb{R}$$

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

Ex $x \sim \text{Geo}(\theta)$, show that $r_x(t) = \frac{\theta}{1-(1-\theta)t}$

$$r_x(t) = \frac{\theta}{1-(1-\theta)e^t}$$

Ex $x \sim \text{Exp}(\lambda)$

$$\begin{aligned} m_x(t) &= E(e^{tx}) = \int_0^\infty e^{tx} f(x) dx \\ &\equiv \int_0^\infty \lambda e^{-\lambda x} e^{tx} dx \\ &= \lambda \int_0^\infty e^{-(\lambda-t)x} dx \\ &= \frac{\lambda}{\lambda-t}, \quad t < \lambda \end{aligned}$$

since it's a cont
var we aren't as interested
in prob generating but
rather moment generating
function

Ex $Z \sim N(0, 1)$ Find moment generating function of Z

$$\begin{aligned} m_z(t) &= E(e^{tz}) = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2 + tz + t^2 - t^2} dz \\ &= e^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} dz \end{aligned}$$

↑ this is correct
above could have
been seen.

normal distribution
with mean t . If integrate density = 1

$$= e^{\frac{1}{2}t^2}$$

$$\therefore \text{if } z \sim N(0, 1) \Rightarrow m_z(t) = e^{\frac{1}{2}t^2}, t \in \mathbb{R}$$

Result if $y = a + bx$ then $m_y(t) = e^{at} m_x(bt)$

$$\underline{\text{PF}} \quad m_y(t) = E[e^{ty}] = E[e^{t(a+bx)}] = E[e^{at+bt^2x}]$$

$$= \underbrace{e^{at}}_{\text{constant}} \quad E[e^{bt^2x}]$$

$$= e^{at} m_x(bt)$$

$$m_x(t) = E(e^{tx})$$



Ex $x \sim N(\mu, \sigma^2)$. Find m_x ?

we know if $x \sim N(\mu, \sigma^2) \Rightarrow z = \frac{x-\mu}{\sigma} \sim N(0, 1)$

$\Rightarrow x$ has the same distribution
 $x = \mu + \sigma z$

where z is standard normal ($z \sim N(0, 1)$)

$$\begin{aligned} m_x(t) &= e^{ut} m_z(st) \\ &= e^{ut} \cdot e^{\frac{1}{2}s^2t^2} \\ &= e^{\mu t + \frac{1}{2}\sigma^2 t^2} \end{aligned}$$

$$m_z(t) = e^{\frac{1}{2}t^2}$$

Results

$$r_x(t) = E[t^x]$$

if x takes values 0, 1, 2, ...

$$\begin{aligned} E g(x) &= \sum_x g(x) p_x(x) \\ &= g(0)p_x(0) + g(1)p_x(1) + \dots \end{aligned}$$

$$r_x(t) = E[t^x]$$

if x takes values $0, 1, 2, \dots$

then

$$r_x(t) = E[t^x]$$

$$= t^0 p_x(0) + t^1 p_x(1) + t^2 p_x(2) \dots$$

$$\textcircled{1} \quad r_x(0) = p_x(0) = P(x=0)$$

$$\textcircled{2} \quad r_x(t) = 1p_x(1) + 2t p_x(2) + 3t^2 p_x(3) + \dots$$

$$\textcircled{3} \quad r_x''(t) = 2p_x(2) + 3 \cdot 2 \cdot t p_x(3) + \dots$$

$$r_x''(0) = 2(p_x(2)) \Rightarrow p_x(2) = \frac{r_x''(0)}{2!} = P(x=2)$$

In general

$$P(x=k) = p_x(k) = \frac{r_x^{(n)}(0)}{k!}$$

, if you know
probability
generating
function of
our generate
all probabilities

$x \sim \text{bin}(n, \theta)$

$$p_x(t) = (1-\theta + \theta t)^n$$

$$P(x=0) = r_x(0) = (1-\theta)^n$$

$$P(x=1) = n\theta (1-\theta)^{n-1}$$

$$\text{show that } P(x=k) = \binom{n}{k} \theta^k (1-\theta)^{n-k}$$

$$\begin{aligned} & E[g(x)] \\ &= \sum_x g(x) p_x(x) \\ &= g(0)p_x(0) + g(1)p_x(1) \dots \end{aligned}$$

Last hour: Generating function

- p.g f. $r_x(t) = E[t^x]$
- m.g f. $m_x(t) = E[e^{tx}]$
- Examples:

$$x \sim \text{Bin}(n, \theta) \quad m_x(t) = r_x(e^t)$$

$$r_x(t) = (1-\theta + \theta e^t)^n$$

$$m_x(t) = (1-\theta + \theta e^t)^n$$

$$r_x(t) = m_x(\ln t)$$

$$x \sim \text{Geo}(\theta);$$

$$r_x(t) = \frac{\theta}{1-(1-\theta)e^t}$$

$$m_x(t) = \frac{\theta}{1-(1-\theta)e^t}$$

$$x \sim P_0(\lambda);$$

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

$$m_x(t) = e^{\lambda(e^t - 1)} \quad \forall t \in \mathbb{R}$$

$$x \sim \exp(\lambda) \Rightarrow$$

$$m_x(t) = \frac{\lambda}{\lambda - 1} \quad t < \lambda$$

$$x \sim N(\mu, \sigma^2) \Rightarrow$$

$$m_x(t) = e^{\mu + \frac{1}{2}\sigma^2 t^2} \quad \forall t \in \mathbb{R}$$

$$y = a + bx \Rightarrow m_y(t) = e^{at} m_x(bt)$$

if x takes values $0, 1, 2, \dots, k$ $\beta(x=k) = \underbrace{r_x(k)}_{k!}$

Result

• $m_x^{(k)}(0) = E[X^k]$ $\begin{cases} \text{derivative} \\ \text{of } x \end{cases}$ $\begin{cases} \text{Kth moment} \\ \text{of } x \end{cases}$

PF $m_x(t) = E[e^{tx}]$

... \vdots

Recall

$$m_x(t) = E(e^{tx})$$

... \vdots

$$\text{Pf } m_x(t) = E[e^{tx}]$$

$$m_x^{(1)}(t) = E[e^{tx} \cdot x] \Rightarrow m_x^{(1)}(0) = E(x)$$

$$m_x^{(2)}(t) = E[x^2 e^{tx}] \Rightarrow m_x^{(2)}(0) = E(x^2)$$

$$m_x^{(k)}(t) = E[x^k e^{tx}] \Rightarrow m_x^{(k)}(0) = E(x^k)$$

$$\text{Ex } x \sim \exp(\lambda) \Rightarrow m_x(t) = \frac{\lambda}{\lambda - t}$$

$$m_x^{(1)}(t) = \frac{\lambda}{(\lambda - t)^2} \quad m_x^{(2)}(0) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda} = E(x)$$

$$m_x^{(2)}(0) = \frac{2\lambda}{(\lambda - t)^3} \quad m_x^{(2)}(0) = \frac{2}{\lambda^2} = E(x^2)$$

$$v(x) = E(x^2) - E(x)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Results

if x_1 & x_2 are \perp then

$$(a) r_{x_1+x_2}(t) = [r_{x_1}(t)] [r_{x_2}(t)]$$

$$(b) m_{x_1+x_2}(t) = [m_{x_1}(t)] [m_{x_2}(t)]$$

Pf b.

$$\begin{aligned} m_{x_1+x_2}(t) &= E[e^{t(x_1+x_2)}] \\ &= E[e^{tx_1} \cdot e^{tx_2}] \end{aligned}$$

$$\left| \begin{array}{l} \text{Recall} \\ m_x(t) = E[e^{tx}] \end{array} \right.$$

$$\begin{aligned} &\simeq E[e^{tx_1}] E[e^{tx_2}] \quad \text{since } \perp \\ &= m_{x_1}(t) + m_{x_2}(t) \end{aligned}$$



In general

$$x_1 + x_2 + \dots + x_n \perp \Rightarrow M_{\sum x_i}(t) = \prod_{i=1}^n M_{x_i}(t)$$

$$r_{\sum x_i}(t) = \prod_{i=1}^n r_{x_i}(t)$$

recall

$$\prod_{i=1}^n a_i = a_1 + a_2 + \dots + a_n$$

Theorem 3.4.6

Not possible for two r.v. to have same moment generating function.

x and y can have the same mean but it does not imply same distribution

Def if x is a r.v. s.t. $M_x(t) < \infty$ for $t \in (-t_0, t_0)$ for some t_0 , and if y is a random variable s.t

$M_y(t) = M_x(t) \quad \forall t \in (-t_0, t_0)$ then x and y should have the same distribution. $x \stackrel{d}{=} y$

So if r.v. has $M_x(t) = \frac{\lambda}{\lambda - t}$ $\Rightarrow x \sim \text{Exp}(\lambda)$

NOTE! This does not apply for mean!

Example Let x_1, x_2, \dots, x_n are independent and $x_i \sim N(\mu_i, \sigma_i^2)$

Find the distribution of $y = \sum_{i=1}^n x_i$

Find moment generating of y then recognize the distribution

$$M_{x_i}(t) = e^{\mu_i t + \frac{1}{2} \sigma_i^2 t^2}$$

$$M_{\sum x_i}(t) = \prod_{i=1}^n M_{x_i}(t)$$

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

use result

$$\therefore M_{\sum x_i}(t) = \prod_{i=1}^n M_{x_i}(t)$$

$$\begin{aligned}
 &= \prod_{i=1}^n \left(e^{\mu_i t + \frac{1}{2} \sigma_i^2 t^2} \right) \\
 &= \prod_{i=1}^n e^{\sum_{i=1}^n (\mu_i t + \frac{1}{2} \sigma_i^2 t^2)} \\
 &= M_{\sum x_i}(t) = e^{(\sum \mu_i)t + \frac{1}{2} (\sum_{i=1}^n \sigma_i^2) t^2}
 \end{aligned}$$

and
 $m_X(t) = e^{\mu t + \frac{1}{2} \sigma^2 t^2}$

$e^{\alpha_1} e^{\alpha_2} \cdots = \prod e^{\alpha_i}$
 $= e^{\sum \alpha_i}$

By uniqueness of moment generating function

$$\sum_{i=1}^n x_i \sim N(\sum \mu_i, \sum_{i=1}^n \sigma^2)$$

But remember for this to be true it has to be normal

Skip 3.5 and go back to it

3.6: Inequalities

(1) Markov's inequality

If x is a non negative r.v then for any $a > 0$

$$P(x > a) \leq \frac{E(x)}{a}$$

Example

$x \sim \text{Exp}(\lambda=3)$. Find the Markov's upper bound for $P(x \geq \frac{5}{2})$

Answer Markov's upper bound = $E(x) = \frac{1}{\lambda} = \frac{1}{3}$

i.e. $P(x \geq \frac{5}{2}) \leq \frac{2}{5}$

(2) Chebychev's inequality

Let x be a n.n r.v $c + Vx < \infty$ then for $x > 0$

Chesby's inequality

Let x be any r.v s.t $\underbrace{V(x)}_{\text{finite}} < \infty$ then for $a > 0$

$$P(|x - \mu_x| \geq a) \leq \frac{V(x)}{a^2}$$

If let $y = (x - \mu_x)^2$ then y is a non neg r.v because of the $(x)^2$, so the markov inequality applies.

$$P(y \geq a^2) = \frac{E(y)}{a^2}$$

could
be
anything

$$\begin{aligned} P((x - \mu_x)^2 \geq a^2) &\leq \frac{E(x - \mu_x)^2}{a^2} \\ &= P(|x - \mu_x| \geq a) \leq \frac{V(x)}{a^2} \end{aligned}$$

Ex $x \sim \text{Exp}(\lambda = 3)$. Find the chesby's upperbound for $P(|x - \frac{1}{3}| \geq \frac{1}{2})$

Ans

$$\text{mean} = \frac{1}{3} = \frac{1}{\lambda}$$

$$(\text{chesby upperbound} = \frac{\text{var}(x)}{\left(\frac{1}{2}\right)^2})$$

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

• Cauchy-Schwarz inequality

if x & y are only two r.v's

$$(E(xy))^2 \leq E(x^2) \cdot E(y^2)$$

Application

if you take $x = (x - \mu_x)$, $y = (y - \mu_y)$

then

$$\begin{aligned} (E((x - \mu_x)(y - \mu_y)))^2 &\leq E[(x - \mu_x)^2] E[(y - \mu_y)^2] \\ \Rightarrow (\text{cov}(x, y))^2 &\leq \text{Var}(x) \text{Var}(y) \end{aligned}$$

$$\Rightarrow \frac{(\text{cov}(x, y))^2}{\text{Var}(x) \text{Var}(y)} \leq 1$$

$$\Rightarrow \frac{(\text{cov}(x, y))^2}{\sqrt{\text{Var}(x) \text{Var}(y)}} \leq 1$$

correlation

$$|\text{cov}(x, y)| \leq 1$$

$$\Rightarrow -1 \leq \text{cor}(x, y) \leq +1$$

Chapter 4: Sampling distribution and limits

Definition

(Random sample)

A collection of iid random variables from some distribution is called a random sample from that distribution.

Definition (statistic)

Any function of a random sample is called a statistic.

Example | Let x_1, x_2, \dots, x_n be a random sample

$$\textcircled{1} \quad \bar{x}_n = \frac{x_1 + x_2 + \dots + x_n}{n} = \sum \frac{x_i}{n}$$

$$\textcircled{2} \quad \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}$$

$$\textcircled{3} \quad \max(x_1, \dots, x_n)$$

The distribution of a statistic is called the sampling distribution of that statistic.

Example 1:

x_1, x_2, \dots, x_n is a r.s from $N(\mu, \sigma^2)$

(a) Find the sampling distribution of $T_n = \sum_{i=1}^n x_i$

(b) Find the sampling distribution of $\bar{x}_n = \underbrace{\sum_{i=1}^n x_i}_{n}$

Ans

(a) x_1, \dots, x_n independent $x_i \sim N(\mu_i, \sigma_i^2)$

then $\Rightarrow \sum_{i=1}^n x_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$

$x_1, \dots, x_n \stackrel{iid}{\sim} N(\mu, \sigma^2) \Rightarrow T_n \sum_{i=1}^n x_i \sim N(n\mu, n\sigma^2)$



$$\text{b) } T_n \sim N\left(n\mu, n\sigma^2\right)$$

$$m_{T_n}(t) = e^{n\mu t + \frac{1}{2}n\sigma^2 t^2}$$

$$M_{\bar{x}_n}(t) = e^{n\mu\left(\frac{t}{n}\right) + \frac{1}{2}n\sigma^2\left(\frac{t}{n}\right)^2}$$

$$= e^{\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2}$$

μ
 σ^2

Recall

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

$\Rightarrow m_{\bar{x}_n}(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

$$M_{\bar{X}_n}(t) = e^{n\mu t + \frac{1}{2} n \sigma^2 t^2}$$

$$\Rightarrow \bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$$

$$\begin{aligned} &\Rightarrow m_x(t) = e^{\mu t + \frac{1}{2} \sigma^2 t^2} \\ &y = a + b x \\ &a = 0 \quad b = \frac{1}{n} \\ &m_y(e) = e^{\mu e} m_x(b e) \end{aligned}$$

Convergence in Probability

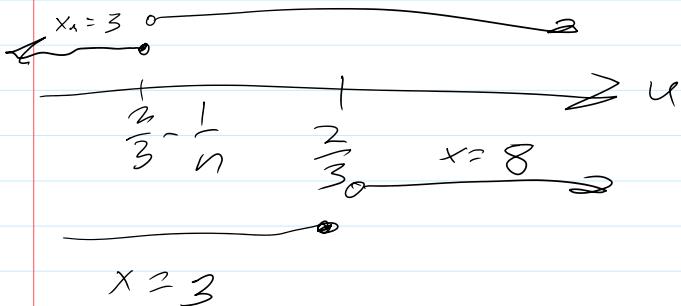
Definition Let x_1, x_2, \dots be a sequence

of random variables then we say
 x_n converges to x in probability if
for any $\epsilon > 0$, $\lim_{n \rightarrow \infty} P(|x_n - x| \geq \epsilon) = 0$

$\forall \epsilon > 0$, $\lim_{n \rightarrow \infty} P(|x_n - x| \geq \epsilon) = 0$ / we write $x_n \xrightarrow{P} x$

$$\text{Ex let } x_n = \begin{cases} 3 & n \leq \frac{2}{3} - \frac{1}{n} \\ 8 & \text{otherwise} \end{cases}$$

Show that $x_n \xrightarrow{P} x$ $x_n = a$



$$u \sim \text{Uni}(0, 1)$$

Let $\epsilon > 0$

$$P(|x_n - x| \geq \epsilon) = P\left(\frac{2}{3} - \frac{1}{n} < u \leq \frac{2}{3}\right)$$

$$= \frac{1}{n}$$

$$0 \leq P(|x_n - x| \geq \varepsilon) = \frac{1}{n} \quad \underline{\text{sequence term}}$$

$$\lim P(|x_n - x| \geq \varepsilon) = 0$$

i.e. $x_n \xrightarrow{P} x$ ~~def~~

Last hour / Sampling distributions (4.1)

Convergence

4.2 Convergence in probability

- Definition:

A sequence of random variables $\{X_n\}$ (i.e. x_1, x_2, \dots) is said to be convergent to x in probability if

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} P(|x_n - x| \geq \varepsilon) = 0$$

Before (in chapter 3)

[Inequalities (3 inequalities)]

- Markov: For any non-neg r.v. X , $P(X \geq a) \leq \frac{E(X)}{a}$
- Chebyshev's: for any r.v. X , $P(|X - \mu_X| \geq a) \leq \frac{\text{Var}(X)}{a^2}$
- Cauchy-Shwartz: for any two random variables $(E(xy))^2 \leq E(x^2)E(y^2)$

Today Convergence in distribution

Weak law of large numbers

Let x_1, x_2, \dots be independent random variables with the same means

$$\Rightarrow E(x_i) = \mu, \forall i$$

and $V(x_i) \leq V \quad \forall i \quad \text{THEN}$

$$M_n = \frac{1}{n} \sum_{i=1}^n x_i \xrightarrow[\text{converges in probability}]{} \mu \quad \left| \begin{array}{l} \text{pf: } E[M_n] = E\left[\frac{1}{n} \sum_{i=1}^n x_i\right] \\ = \frac{1}{n} \sum_{i=1}^n E[x_i] = \frac{1}{n} \cdot n \cdot \mu = \mu \\ \text{due to linearity} \end{array} \right.$$

PF By Chebyshev's inequality

Let $\varepsilon > 0$

$$P[|M_n - \mu| \geq \varepsilon] \leq \frac{V(M_n)}{\varepsilon^2}$$

$$\begin{aligned} V(M_n) &= V\left(\frac{1}{n} \sum_{i=1}^n x_i\right) \\ &\stackrel{\text{prop!}}{=} \frac{1}{n^2} \sum_{i=1}^n V(x_i) \leq \frac{1}{n^2} \sum_{i=1}^n V \\ &= \frac{1}{n^2} \cdot n V \end{aligned}$$

$$\begin{aligned} n \sum_{i=1}^n M_i &= n \bar{M}_n \\ &= \frac{1}{n} \cdot n \bar{M}_n \\ &= \bar{M}_n \end{aligned}$$

$$0 \leq P[|M_n - \mu| \geq \varepsilon] \leq \frac{\text{Var}(M_n)}{\varepsilon^2} \stackrel{n \rightarrow \infty}{\rightarrow} 0$$

by squeeze theorem

$$\lim_{n \rightarrow \infty} P(|M_n - \mu| \geq \varepsilon) = 0$$

$$\text{i.e. } M_n \xrightarrow{P} \mu \quad \boxed{\text{}}$$

Today Convergence in distribution

Definition

Let x_1, x_2, \dots be a sequence of random variables. ($\{x_n\}$) and x is a random variable then we say

$\{x_n\}$ converge to x in distribution if

$$\lim_{n \rightarrow \infty} P(x_n \leq x) = P(x \leq x), \quad \forall x \mid P(x=x) = 0$$

Example

$$X_n \sim \text{Bin}(\theta = \frac{1}{n})$$

		P _X (x)
		x
		0
		1 - $\frac{1}{n}$
		$\frac{1}{n}$

$$\begin{cases} \text{RCC=1} \\ P_n(x) = \begin{cases} \frac{1}{n}, & x < 0 \\ 1 - \frac{1}{n}, & 0 \leq x < 1 \\ \frac{1}{n}, & x \geq 1 \end{cases} \end{cases}$$

$$\lim_{n \rightarrow \infty} P(x_n \leq x) = P(x \leq x) \iff \lim_{n \rightarrow \infty} F(x) = F(x)$$

$$F_{x_n}(x) = \begin{cases} 0, & x < 0 \\ 1 - \frac{1}{n}, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

$$x = \begin{array}{|c|} \hline 1 \\ \hline \end{array} \quad F_x(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

Show that $x_n \xrightarrow{D} x$ where x_n converge to x in distribution

$$F_x(x) = \begin{cases} 0 & , x < 0 \\ 1 & , x \geq 0 \end{cases}$$

$$\text{pf } \lim_{n \rightarrow \infty} F_{x_n}(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

$$= F_x(x), \forall x$$

$$\therefore x_n \xrightarrow{D} x$$

Central Limit Theorem

Let x_1, x_2, \dots be iid r.v's with finite mean $E(x_i) = \mu$ and finite variance. $\text{Var}(x_i) = \sigma^2$, then

$$Z_n = \frac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{D} Z \text{ where } Z \text{ is standard normal distribution}$$

$$M_n = \frac{1}{n} \sum_{i=1}^n x_i$$

Example

x_1, x_2, \dots iid with $E(x_i) = \mu = 67$ and $\text{Var}(x_i) = \sigma^2 = 10^2$

Find the approx probability $P(M_{100} \leq 70)$

$$Z_{100} = \frac{M_{100} - \mu}{\frac{\sigma}{\sqrt{100}}} \xrightarrow{D} Z \quad (\text{CLT})$$

$$= \frac{M_{100} - 67}{\frac{10}{\sqrt{100}}} \xrightarrow{D} Z$$

$$\left(\frac{10}{10} \right) = 1$$

$$P(M_{100} \leq 70) = P(M_{100} - 67 \leq 70 - 67)$$

$$\left(\frac{1}{1} \right)$$

Closely $\rightarrow n \sim N(67, 10^2)$

~~Closed
not exact~~ $\rightarrow \text{xp}(\mathbb{P}(Z \leq 3) = 1 - \mathbb{P}(Z \leq -3))$
 $= 0.9987$

use the table

~~4.2~~ convergence in prob
~~4.3~~ Conv S \leftarrow stopped it, not on mean and 3.5
4.4 convergence in Distribution