May 12, 2018 12:34 PM

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:

Axiom ①: $0 \leq P(A) \leq 1$
Axiom ②: $P(S) = 1$
Axiom ③: if $A_1, A_2 \ldots$ is a collection of disjoint events the $P(A_1 \cup A_2 \cup A_3 \ldots)$
$= P(A_1) + P(A_2) + P(A_3) \ldots = P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

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

Result ①: $P(\emptyset) = 0$

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

Pf/ Let $A_i = \emptyset$
$i = n+1, n+2 + \ldots$

Then $A_i \cap A_j = \emptyset$ $\forall i \neq j$

By Axiom ③ the $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

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

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

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

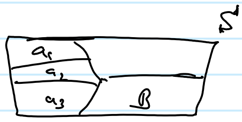
In particular, if A & B are disjoint then $P(A \cup B) = P(A) + P(B)$

Result: Therom of total probability ③

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

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

and let B be any event in S

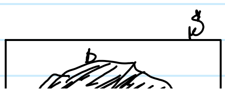


THEN:

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

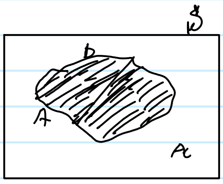
* This can be finite as well

Pf/



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

Pf/



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

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

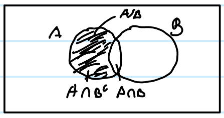
& $B\cap A_1, B\cap A_2 \ldots$ are disjoint

$P(B) = P(B\cap A_1)\cup P(B\cap A_2)\cup\ldots$

$= P(B\cap A_1) + P(B\cap A_2) + \ldots$ ($\because$ $B\cap A_i$'s are disjoint)

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

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

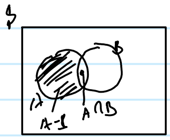


Result ④:

For any two events in S

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

Pf/



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

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

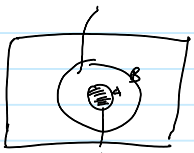
$= P(A\setminus B) + P(AB)\ldots$

$P(A\setminus B) = P(A) - P(AB)$ ($\because$ $A\setminus B$ and AB are disjoint)

Result ⑤:

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

$B\cap A^c$ or B/A



$A\cap B = AB$

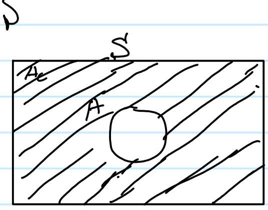
Pf/ $P(B\setminus A) = P(B) - P(AB)$

$= P(B) - P(A) \geq 0$

$\Rightarrow P(A)\leq P(B)$ ■

Result 6: For any event in S

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



Pf/ $A^c = S - A$

$P(A^c) = P(S\setminus A)$

$= P(S) - P(SA)$

$= 1 - P(A)$ ⟵ by axiom ②

Result 7

For any two events A & B in 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

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

$\blacksquare$

Lecture 01

Probability

Random Experiments: Experiments when the event is random

Let A be an event in a random experiment

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

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

Events

Events are a subset of a sample space

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

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

$\{\}$ = empty set
$\emptyset$ = not set
$\{\} == \emptyset$

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

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

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

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

* Note don't repeat elements in sets

$\{1,2,3,3,4\}$ ✗ $\{1,2,3,4\}$ ✓

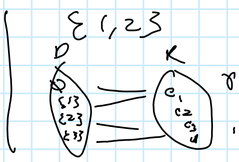
Definition of Probability

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

Axis ① $0 \le P(A) < 1$

② $P(S) = 1$

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



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

① $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

② $(A \cap B)^c = A^c \cup B^c$

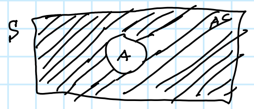
③ $(A \cup B)^c = A^c \cap B^c$

④ $A \cup \emptyset = A$

⑤ $A \cap \emptyset = \emptyset$

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

A^c = compliment



$A \cap B == AB$

Last hour

Counting: Given the digits $\overbrace{[1][2]\ldots[9]}^{n}$

how many 4 digit numbers can be formed?

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

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

Combinations: The number of combination of k object 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 repitition: n^r
Permutation without repitition: $\frac{n!}{(n-r)!}$
Combination with repitition: $\frac{(r+n-1)!}{r!(n-1)!}$
Combination without repitition: $\frac{n!}{(n-r)!r!}$

The number of ways of dividing n objects into groups of sizes
$n_1\, n_2 \ldots n_k$ is: $\frac{n!}{n_1!\,n_2!\ldots n_k!} = \binom{n}{n_1\, n_2 \ldots n_k}$

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

$\overline{10}\ \overline{9}\ \overline{8}\ \overline{7}\ \overline{6}$

$\frac{n!}{(n-r)!} = \frac{10!}{(10-5)!} = \frac{10!}{5!}$ → # of possibilities in the exam

$\frac{\binom{6}{5} \times \binom{4}{0}}{\binom{10}{5}}$

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

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

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

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

$$= \frac{\frac{12!}{3!3!3!3!} \times \frac{40!}{10!10!10!10!}}{\frac{52!}{13!13!13!13!}}$$

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?

$$\frac{60!}{30!\,30!} \qquad = \frac{\binom{58}{28}+\binom{58}{28}}{\binom{60}{30}}$$

Conditional Probability (1.5)

Def^n 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, $P(A)=0.5$, $P(B)=0.3$, $P(AB)=0.1$

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

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

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

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

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

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

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 S s.t $P(B) > 0$

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

axioms

$p(A|B) \leq 1$
$p(S|B) = 1$
$p(\cup A_i|B) = \sum_{i=1}^{n} P(A_i|B)$

Note: conditional probability is a probability measure.

Pf $A \cap B \subseteq B$
$0 \leq P(A \cap B) \leq P(B) = 1$

$0 \leq P(A|B) \leq 1$ Axiom 1

Axiom 2

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

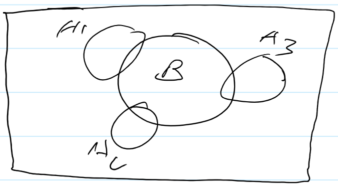
$P(B)\underline{P(B)}$

$= 1$

Axiom 3

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

$$P(A_1 \cup A_2 \cup \ldots | B) = \frac{P((A_1 \cup A_2 \ldots) \cap B)}{P(B)} \ldots \text{(Def}^n\text{)}$$



$$= \frac{P(A_1 B \cup A_2 \cap B, A_3 \cap B \ldots)}{P(B)}$$

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

($A_1 B_1$, $A_2 B_2$ are disjoint)

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

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

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

The Theroem of Total Probability

Conditional Probability version.

Let $A_1, A_2, \ldots$ be a partition of S

partion means:

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

and B be any event in S Then

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

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

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

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

Theroem Bayes Theroem

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

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

Pf

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

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

$$P(AB) = P(B|A)P(A)$$

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

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

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 found to favor the issue. Find the conditional probability that he is a democrat

$P(r) = 0.4$
$P(d) = 0.6$
$P(S) = P(r \cup D) = 1$

$P(F \mid R) = 0.3$
$P(F \mid D) = 0.5$

① $P(D \mid F) = \dfrac{\overset{0.6}{P(D)}}{P(F)} \times \overset{0.5}{P(F \mid D)}$

② $P(F) = P(F \mid R)P(\overset{A_1}{R}) + P(F \mid D)P(\overset{A_2}{D})$

$= 0.3(0.4) + 0.5(0.6)$

$= 0.12 + 0.3$

$= 0.42$

③ $P(D \mid F) = \dfrac{0.6}{0.42} \times 0.5$

1.5.2 Independent events

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

Note: if $P(B) > 0$

if $P(B) > 0$

$A \perp B$ iif $P(A \mid B) = P(A)$

independent of

$P(A \mid B) = \dfrac{P(AB)}{P(B)}$

if independent

$P(A \mid B) = \dfrac{P(A) \times \cancel{P(B)}}{\cancel{P(B)}}$

$P(A \mid B) = P(A)$

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

A = observe an odd number = $\{1, 3, 5\}$
B = Observe an even number = $\{2, 4, 6\}$
C = observe a number $< 3 = \{1, 2\}$

Q. a) Are A & B independent

$P(A) = \frac{1}{2}$ $\quad AB = \emptyset \quad$ $P(AB) = P(A) \times P(B)$

$P(B) = \frac{1}{2}$ $\quad P(AB) = \emptyset \quad$ $0 = 0.5 \times 0.5$

$P(C) = \frac{1}{3}$ $\quad 0 \neq 0.25$

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

Pf: $P(A \cap B^c) = P(A \setminus B)$

$= P(A) - P(AB)$

$= P(A) - P(A)P(B)$

$= P(A)[1 - P(B)]$

$= P(A)P(B^c)$

$\therefore A \perp B^c$ $\blacksquare$

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

To prove:

$P(AB^c) = P(A)P(B^c)$

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 $P(A_{i_1} \cap A_{i_2} \cap \dots A_{i_k}) = P(A_{i_1}) \times P(A_{i_2}) \cdots P(A_{i_k})$ for each subcollection of events $A_{i_1}, A_{i_2}, \dots A_{i_k}$

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

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

$\boxtimes$

Ex/ Roll a 4 sided (tetrahedral) die.

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

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

(a) are A & B independent?

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

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

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

$\therefore A \perp B$

$P(A) = P(B) = P(C) = \frac{2}{4} = \frac{1}{2}$

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

$P(BC) = \{1\}$

(b) Are $B \perp C$?

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

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

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

(c) Are $A \perp C$

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

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

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

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 $A_1, A_2 \ldots$ be a collection of sets

We say A_n increases A ($A_n \uparrow A$) if $A_1 \subseteq A_2 \subseteq A_3 \ldots$

and $A = \bigcup_{i=1}^{\infty} A_i$

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

f is cts at a if $\lim_{x \to a} f(x) = f(a)$

Recall

F is continuous at a if for every sequence $\{x_n\}$ s.t. $\lim_{n \to \infty} x_n = a$

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

Definition: we can $\{A_n\} \downarrow A$ if $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$

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

and

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

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

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

$f\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} (x_n)$

Theorem: 1.6.1 Continuity of Probability

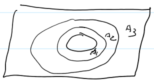
If $\{A_n\} \uparrow A$ or $\downarrow A$, then

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

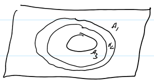
May 22, 2018 12:03 PM

Last hour

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



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



(we call this $\lim_{n\to\infty} A_n$)

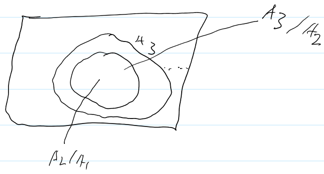
Therom 1.6.1 Continuity therom of probability

if $\{A_n\} \uparrow A$ or $\{A_n\} \downarrow A$ then $P(A) = \lim_{n\to\infty} P(A_n)$

Proof case 1:

$\{A_i\} \uparrow A$

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



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

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

$\because A_1 = \bigcup_{i=2}^{\infty} (A_2 \setminus A_{i-1})$ 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$)$

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

* Remember: $P(A \setminus B) = P(A) - P(AB)$

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

* $P(A_i A_{i-1}) = P(A_i)$ because $A_i \subseteq A_{i-1}$

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

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

$$= P(A_n) - P(A_1)$$

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

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

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

Case 2 If $\{A_n\} \downarrow A = \bigcap_{n=1}^{\infty} A_n$

Let $B_n = A_n^c$ then

$$B_n \uparrow \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n^c = \underbrace{\left(\bigcap_{n=1}^{\infty} A_n\right)^c}_{= A^c}$$ ← De morgans law

i.e $\{B_n\} \uparrow A^c$

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

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

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

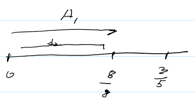
$$1 - P(A) = \lim_{n\to\infty} (1 - P(A_n))$$

$$X - P(A) = X - \lim_{n\to\infty} P(A_n)$$

$$P(A) = \lim_{n\to\infty} P(A_n) \qquad \blacksquare$$

<u>Example</u>: 1.6.4 p30/ Suppose $P\Big(\big[\overbrace{0, \tfrac{8}{4+n}}^{A_n}\big]\Big) = \dfrac{2 + e^{-n}}{6}$, $i = 1, 2, \ldots$

Find $P(\{0\})$?



Solution: $A_n \downarrow \{0\}$!

Solution:

$$P(\{0\}) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty}\left(\frac{2+e^{-n}}{6}\right) = \frac{2}{6} = \frac{1}{3}$$

Ex/ $S = \{1,2,3,\ldots\}$ $\quad P(\{s\}) = \left(\frac{1}{2}\right)^s$

Find $P(\{5,6,7,\ldots\})$ using continuity of probability and finite additivity

<u>Solution</u> Define $A_n = \{5,6,\ldots,n\}$ then $A_n \uparrow A = \{5,6,\ldots\}$
using continuity of probability. $P(\{5,6,\ldots\}) = \lim_{n\to\infty} P(A_n)$

$A_n = \{5, 6 \ldots n\} \leftarrow$ n-4 elements in A

<u>Geometric sequence/series</u>

$$S_n = \frac{a(1-r^n)}{1-r}$$

$$P(A_n) = P(\{5\}) + P(\{6\}) + \ldots + P(\{n\})$$

$$= \underbrace{\left(\tfrac{1}{2}\right)^5}_{a} + \underbrace{\left(\tfrac{1}{2}\right)^6}_{r=\frac{1}{2}} + \ldots \left(\tfrac{1}{2}\right)^n$$

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

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

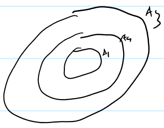
$$\lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} \left(\tfrac{1}{2}\right)^4\left[1-\left(\tfrac{1}{2}\right)^{n-4}\right] = \left(\tfrac{1}{2}\right)^4 = \frac{1}{16}$$

May 24, 2018 12:02 PM

Last hour (continuity of probability)

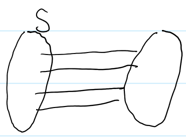
Midterm * tentative: July 6^{th}

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



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

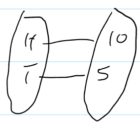
Definition: A random variable (r.v) is a function from the sample space to the set of real numbers ($\mathbb{R}$).



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

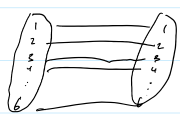
$x(H) = \$10$

$x(T) = \$-5$



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) = $ # of heads in S

$x(HH) = 2$ $\quad$ $x(TH) = 1$

$x(HT) = 1$ $\quad$ $x(TT) = 0$

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

$$x(s) = \begin{cases} 1 & s \in \{5, 6\} \\ 0 & o/w \end{cases}$$

Indicator Function

Indicator Function

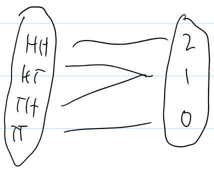
$$I_A(s) \begin{cases} 1 & s \in A \\ 0 & s \notin A \end{cases}$$

Definition

$X = Y$ means $X(s) = Y(s)$ $\forall s \in S$

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

Find $X \in \{1, 2\}$



Ans $= \{HH, HT, TH\}$

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

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

$X(\text{cl}) = 200$
$X(\text{sn}) = 100$
$X(R) = -50$

Find the distribution of X?

Let $B \subseteq \mathbb{R}$

$B = (190, 210)$
$P(X \in (190, 210)) = P(X = 200)$

Let $B \subseteq R$

$P(x \in B) = 0.5\, I_B^{(200)} + 0.2\, I_B^{(100)} + 0.3\, I_B^{(-50)}$

$B = (190, 210)$

$P(x \in (190, 210)) = P(x = 200)$
$= P(1)$
$= 0.5$

$B = \{90, 210\}$

$P(x \in (90, 210)) = P(x = 100 \text{ or } x = 200)$
$= 0.2 + 0.5$
$= 0.7$

$P(x \in (-60, 500)) = 1$

Ex/ Toss a coin twice $S = \{\overset{\frac{1}{4}}{HH}, \overset{\frac{1}{4}}{HT}, \overset{\frac{1}{4}}{TH}, \overset{\frac{1}{4}}{TT}\}$

you win

\$4 if you get 2 H's
\$1 if you get 1 H
-\$5 if you get no heads

Let x = amount of money you win. Find the distribution of x.

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

Discrete Random Variables

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

$p(x = x_i) = p_i \quad i = 1, 2$ and $\sum_i p_i = 1$

Definition: Probability mass function of discrete random variables.
The prob mass function (pmf) of a discrete random variable

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

Ex

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

$=1$ 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

$=1$

$$P(X \in (4.5, 6.5)) = \frac{1}{6} + \frac{1}{6} =$$

$$= P_X(5) + P_X(6)$$

$$= \sum P_X(x) \quad x \in (4.5, 6.5)$$

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

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	$\frac{1}{2}$					$\frac{1}{2}$

1

② Bernoull distribution

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

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

X	0	1

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

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 & θ $(x \sim \text{Binomial}(n,\theta))$ if $P_x(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$, $x = 0, 1, 2 \ldots n$

(=0 o/w)

$x = \#$ 1's

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

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

n	0, ... , n
P_x	

1

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

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

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

4 Geometric Distribution

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

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

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

x = # of failures before the first success

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

x	0 1 . . . .
P_x	

1

Q^x: $\sum_{n=0}^{\infty} P_x^{(x)} = 1$?

$\sum^{\infty} p(x) \quad \sum^{\infty} \quad x \quad = \sum^{\infty} \quad x$

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

$$= \theta \left[1 + (1-\theta) + (1-\theta)^2 + (1-\theta)^3 + \dots \right]$$

$\Rightarrow$ Geometric series
$a = 1$
$r = (1-\theta)$

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

$\therefore$ $P_x(x)$ is pmf.

(5) Negative Binomial distribution:

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

X = # T's before the r^{th} success

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

$y = 0, 1, 2 \dots$

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

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

May 29, 2018 12:03 PM

Midterm Test: July 6 (19:00 - 21:00) → (7:00 - 9:00 pm)

Rooms AA112, 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): $P_X(x) = \begin{cases} 1-\theta & x=0 \\ \theta & x=1 \end{cases}$

x	0	1
P_X	$1-\theta$	θ

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

x	0 ... n
P_X	. . . .

$= 1$

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

# of failures before the first success → 1 success

⑤ Negative binomial distribution → $1 \leq$ success

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

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

Q_1 $\sum \binom{x+r-1}{x} (1-\theta)^x \theta^r = ?$

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

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

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

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

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

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

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

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

$$= \frac{(-1)^x (-r)(-r-1)(-r-2)\cdots(-r-(x-1))}{1\cdot 2\cdots x}$$

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

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

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

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

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

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

$\therefore$ P_X is a pmf

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

Let x be the # of failures before the third success ; $\sim$ Neg binomial

$r = 3$

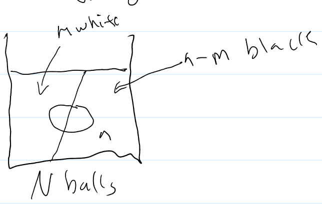
$\theta = 0.6$

$$P(X \geq 2) = 1 - P(X=1) - P(X=0) \qquad \Big|\ P(X=x) = \binom{x+r-1}{x}(1-\theta)^x \theta^r$$

$$= 1 -$$

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

(6) Hypergeometric distribution



$$P_X(x) = P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

A random variable x is said to have a hypogeometric distribution if

A random variable x is said to have a hypogeometric distribution if ($x \sim$ hypogeometric (N, M, n)) ! s

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

$$\left.\begin{array}{l} ① \ x \leq M \\ ② \ x \leq n \end{array}\right\} \quad x \leq \min(n, M)$$

③ $x \geq 0$

④ $n - x \leq N - M$

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

⑦ Poisson distribution

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

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

recall

$e^{\lambda} = 1 + \dfrac{\lambda}{1!} + \dfrac{\lambda^2}{2!} \ldots \ldots$

Term Test: July 6^{th}: 19:00–21:00
Rooms: AA112, AC223

Last hour: Important discrete distributions

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

② Bernoulli (θ) distribution $P_X(x) = \theta^x (1-\theta)^{1-x}$, $x = 0$ or 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 1^{st} success)

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

⑤ Negative binomial distribution [number of failure before r^{th} success]

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

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

$P_X(x) = \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$, $\max(0, n-(N-M)) \leq x \leq \min(n, M)$

⑦ Poisson distribution ($X \sim \text{Poisson}(\lambda)$) $\Leftarrow$ Questions for Poisson will be told but not for the others

$P_X(x) = \dfrac{e^{-\lambda} \lambda^x}{x!}$ $x = 0, 1, 2, \dots$

Poisson 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 of time. $\Longleftarrow$ * exam question

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 [$t = 2$]

$X \sim P_o(\lambda t = 2 \times 2 = 4)$, $P(X = 5) = \dfrac{e^{-4} 4^5}{5!}$

b) what is the probability no calls during the next 10 minutes

X = # of calls in the next 10 min $\sim P_o(\lambda t) = P_o(20)$
$2\times 10 = 20$

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

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 $P(X=x)=0\ \forall x\in\mathbb{R}$

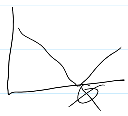
if $P(X=x)=0 \quad \forall x\in\mathbb{R}$

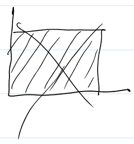
Definition

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

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

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

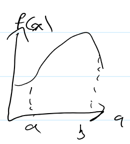




Definition: Absolutely continuous random variables

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

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



Note: $P(X=x) = P(x\leq X\leq x) = \int_x^x f_X(x)\,dx = 0$

Note: A random variable with Absolute continuous $\Rightarrow$ continuous

Note converse is false (cts $\not\Rightarrow$ abs cts)

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

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

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

(a) Find the value of k

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

(a) $\int_{-\infty}^{\infty} f(x)dx = 1$ is $\underbrace{\int_{-\infty}^{1} f(x)\,dx}_{0} + \int_{1}^{5} \underbrace{f(x)}_{k*x} dx + \int_{5}^{\infty} \underbrace{f(x)\,dx}_{0} = 1$

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

$k\left[\frac{25}{2} - \frac{1}{2}\right] = 1$
$k[12] = 1$
$k = \frac{1}{12}$

(b) $P(2 \le x \le 3) = \int_{2}^{3} f(x)\,dx$

$= \int_{2}^{3} \frac{1}{12} x\,dx = \frac{1}{12}\left[\frac{x^2}{2}\right]^3$

$= \frac{1}{12}\left[\frac{9}{2} - \frac{4}{2}\right] - \frac{1}{12} \times \frac{5}{2} = \frac{5}{24}$

Ex2 A random variable x has pdf (probability density function)

$$f_x(x) = \begin{cases} 0.2 & -1 < x \le 0 \\ 0.2 + cx & 0 < x \le 1 \\ 0 & o/w \end{cases}$$

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

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

$\underbrace{\int_{-\infty}^{-1} f(x)dx}_{0} + \int_{-1}^{0} \underbrace{f(x)}_{0.2}dx + \int_{0}^{1} \underbrace{f(x)}_{0.2+cx}dx + \int_{1}^{\infty} \underbrace{f(x)dx}_{0} = 1$

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

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

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

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

$$c = 1.2$$

(b) $P(-0<x<0.5) = \int_{-0.5}^{0.5} f(x)dx$

$$= \int_{-0.5}^{0} \underbrace{f(x)}_{0.2} dx + \int_0^{0.5} \underbrace{f(x)}_{0.2+1.2x} dx$$

$$= \left[0.2x\right]_0^0 + \left[0.2x + \cancel{1.2}^{0.6} \frac{x^2}{\cancel{2}}\right]_0^{0.5}$$

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

Important absolute continuous distribution

① Uniform distribution

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

$$f_X(x) = \begin{cases} \frac{1}{R-6}, & L \le x \le R \\ 0, & o/w \end{cases}$$

eyglx1lnquiy6lsgdk7pgetdl

Note: $\int_{-\infty}^{\infty} f(x)\,dx = \underbrace{\int_{-\infty}^{L} f(x)\,dx}_{0} + \underbrace{\int_L^R f(x)\,dx}_{\frac{1}{R-L}} + \underbrace{\int_R^{\infty} f(x)\,dx}_{0}$

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

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

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

find density function

$$f_X(x) = \frac{1}{6-5} = \frac{1}{11} = 0 \qquad -5 \le x \le 6 \quad \text{o/w} \qquad \text{before } -5 = 0$$

$$P(1 < x < 3) = \int_1^3 f(x)\,dx = \int_1^3 \frac{1}{11}\,dx = \frac{1}{11}x^3 = \frac{1}{11}(3-1) = \frac{2}{11}$$

b) $P(x < -1)$

815570290

From <https://osap.gov.on.ca/OSAPSecurityWeb/public/retrieve_oan_confirmation.xhtml>

$$= \int_{-\infty}^{-1} f(x)\,dx = \underbrace{\int_{-\infty}^{-5} f(x)\,dx}_{0} + \underbrace{\int_{-5}^{-1} f(x)\,dx}_{\frac{1}{11}} = \int_{-5}^{-1} \frac{1}{11}\,dx = \frac{1}{11}(x)\Big|_{-5}^{-1} = \frac{4}{11}$$

c) $P(|x| \le 0.25) = P(-2.5 < x < 2.5)$ $\qquad |x| < a \iff -a < x < a$

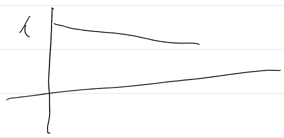
$$= \int_{-2.5}^{2.5} f(x)\,dx$$

$$= \pi\,[x]_{-0.25}^{0.25} = \frac{0.5}{11}$$

② Experimental Distribution:

A random variable x is said to have an exponential distribution with parameters λ ($\lambda > 0$) if its pdf is

$$f_X(x) = \lambda e^{-\lambda x} = 0 \qquad , x \ge 0 \quad \text{o/w}$$



Note: is $f(x) = \lambda e^{-\lambda x}$, $x \ge 0$ a pdf?

$$\int_{-\infty}^{\infty} f(x)\,dx = \underbrace{\int_{-\infty}^{0} f(x)\,dx}_{0} + \int_0^{\infty} f(x)\,dx = \int_0^{\infty} \lambda e^{-\lambda x}\,dx = -\left[-e^{-\lambda x}\right]_0^{\infty} = 1(0) = 1$$

Ex/ $x \sim$ exponential $(\lambda = 2)$

find $P(X > 10)$

$f_X(x) = \lambda e^{-\lambda x} = 0.2 e^{-0.2x}$

Ans $\int_{10}^{\infty} f(x)\,dx$

$= \int_{10}^{\infty} f(x)\,dx = \int_{10}^{\infty} 0.2 e^{-0.2x}\,dx$

$= \left[-e^{-0.2x} \right]_{10}^{\infty}$

$= e^{-0.2 \times 10}$

$= e^{-2}$

June 11, 2018 9:12 PM

Term Test: July 6^{th}, 19:00-21:00

Room AA112, AC223 → Allowed cheatsheet 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

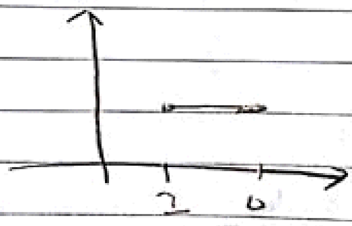
A random variable is absolute continuous if $\exists$ a pdf f_X s.t $P(a \le X \le b) = \int_a^b f_X(x)\,dx$

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 $L \le x \le R$

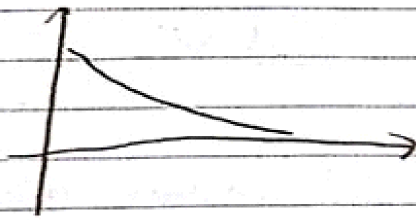
$= 0$ /w



② Exponential distribution

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

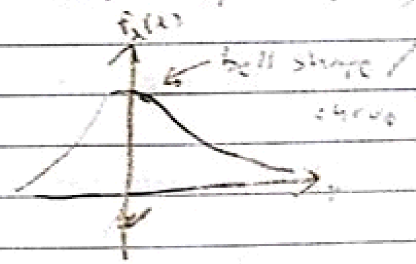
$f_x(x) = \lambda e^{-\lambda x}$ if $x > 0$
$= 0$ o/w



Today: Normal distribution

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

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



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

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

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

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

- PDF 2 conditions ≥ 0

Condition ②:

$$\iint f(x,y)\,dx\,dy = \iint f(r\cos\theta, r\sin\theta)\,r\,d\theta\,dr$$ Recall, integral areas/volume

 $\int_a^b f(x)\,dx$

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

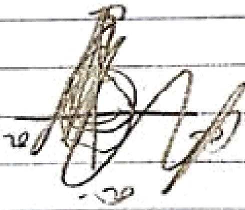
$$\Longleftrightarrow \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}\,dx = \sqrt{2\pi}$$

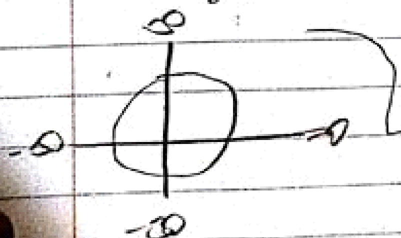
Let $I = \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}\,dx$ then $I = \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\,dy$

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

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

$$= \int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}(\underbrace{r^2\cos^2\theta + r^2\sin^2\theta}_{1 \text{ factor out } r^2})}\,r\,dr\,d\theta \to e^{-\frac{1}{2}r^2}$$





$$= \int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}r^2}\,r\,dr\,d\theta$$
$0-(1)$

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

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

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

June 7, 2018 12:10 PM

Term Test: July 6^{th} 11:00-21:00 (1 hr 50 min)

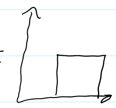
psya

Rooms: AA 112, AC223

Last hour: Important absolute continuous distributions

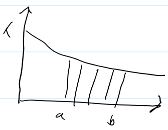
① Uniform $[L, R]$ distributions ($L \leq R$)

$$f_X(x) = \frac{1}{R-L} \quad L \leq x \leq R$$
$$= 0 \quad o/w$$



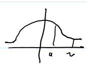
② Exponential Distributions ($\lambda > 0$)

$$f_X(x) = \lambda e^{-\lambda x} \quad x \geq 0$$
$$= 0 \quad o/w$$



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

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



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

④ Gamma (α, λ) distribution ($\alpha > 0, \lambda > 0$)

Recall $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, dx \longleftarrow$ gamma function

Some properties includes:

① $\Gamma(1) = \underline{1}$

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

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

④ $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

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

$= \int_0^\infty \cancel{x}^0 e^{-1} \, dx$

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

$= 1$

Gamma distribution

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

so.

$\int_0^\infty t^{\alpha-1} e^{-t}$, sub $t = \lambda x$, $\lambda > 0$?

$\int_0^\infty t^{\alpha-1} e^{-t}$, sub $t = \lambda x$, $\lambda > 0$ $\Big\}$ u-sub
$dt = \lambda\, dx$

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

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

1

$$\boxed{f_X(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \text{ is a PDF}}$$

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

$= 0$ o/w

$(X \sim \text{Gamma}(\alpha, \lambda))$

Important Note:

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

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

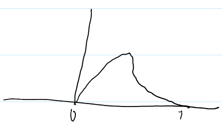
Note Gamma$(1, \lambda)$ = exponential λ

⑤ Beta Distribution

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

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

$= 0$ o/w



$\int_0^\infty \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} dx = 1 \Rightarrow \int_0^\infty x^{\alpha-1} e^{-\lambda x} = \frac{\Gamma(\alpha)}{\lambda^\alpha}$ / $\int_0^\infty x^{9} e^{-\lambda x} dx$, a

$= \frac{\Gamma(10)}{\lambda^{10}}$

$$\int_0^1 x^{10}(1-x)^7\,dx$$

$=$

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

$= 0$ o/w

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

2.5 Cumulative distribution function (cdf)

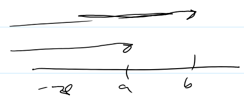
Definition (cdf): Given a random variable X, its cumulative distribution function is the function $F_X : \mathbb{R} \to [0,1]$ defined by $F_X(x) = P(X \le x)$

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

Properties of cdf (F_X)

① $0 \le F_X(x) \le 1 \quad \forall x \in \mathbb{R}$

② $a < b \Rightarrow F_X(a) \le F_X(b)$



Pf $a \le b \Rightarrow X \in (-\infty, a] \overset{\text{subset}}{\subseteq} X \in (-\infty, b]$

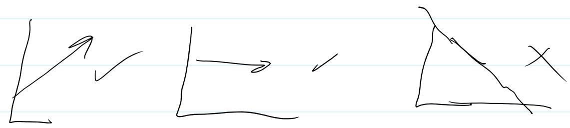
$\Rightarrow P(X \in (-\infty, a]) \le P(X \in (-\infty, b])$

$\iff F_X(a) \le F_X(b)$

$\Rightarrow F_X(a) \le F_X(b)$ ▨

if $a < b \Rightarrow f(a) \le f(b)$

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



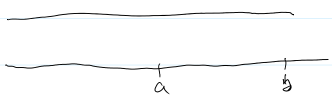
③ $\lim_{x \to \infty} F_X(x) = 1$

$F_X(x) = P(X \le x)$

④ $\lim_{x \to -\infty} F_X(x) = 0$

⑤ for $a < b$, $P(a < X \le b) = F_X(b) - F_X(a)$

Pf:



$P(a < X \le b) = P(X \in (a, b])$

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

$P(X\in(a,b]) = P(X\in(-\infty,b] \setminus X\in(-\infty,a])$, Recall: $P(A\setminus B) = P(A) - P(AB)$

$= P(X\in(-\infty,b]) - P(X\in(-\infty,a])$ ... $(X\in(-\infty,a]$ & $X\in(-\infty,b])$ if $B\subset A$

$AB = B$

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

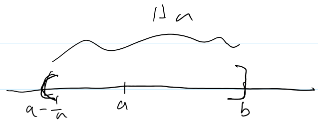
$= F(b) - F(a)$ ■

Property 6

(6) $P(a\le X\le b) = P(X\in[a,b]) = F_X(b) - F_X(a-)$

where $F_X(a-) = \lim\limits_{n\to\infty} F_X(a-\frac{1}{n})$

Proof:



$X = (a-\frac{1}{n}, b] \downarrow X\in\bigcap_{n=1}^{\infty}(a-\frac{1}{n}, b] = X\in[a,b]$

$a \in A_n = (a-\frac{1}{n}, b] \quad \forall n\in\mathbb{N}$

i.e.

$a\in\bigcap_{n=1}^{\infty} A_n$

Using continuity of probability:

$P(X\in[a,b]) = \lim\limits_{n\to\infty} P(X\in(a-\frac{1}{n},b])$

$= \lim\limits_{n\to\infty}\left[F_X(b) - F_X(a-\frac{1}{n})\right]$

$= \lim\limits_{n\to\infty} F_X(b) - \lim\limits_{n\to\infty} F_X(a-\frac{1}{n})$

$= F_X(b) - F_X(a-)$ ■

(7) $P(X\in(a,b)) = F_X(b-) - F_X(a)$ (exercise prove it)

(8) $P(X\in[a,b)) = F_X(b-) - F_X(a-)$

CDF's of discrete random variables

Ex Bernoulli(θ)

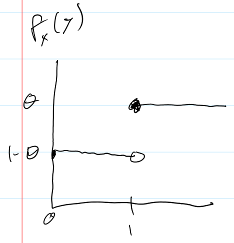
x	p_x

$X \overset{\text{is}}{\sim} Ber(\theta)$

x	p_x
0	$1-\theta$
1	θ

$X \overset{\text{is}}{\sim} Ber(\theta)$

$F_X(-0.5) = P(X \le -0.5)$
$= 0$ ← since $Ber(\theta)$ can only be 0, 1



$F_X(0) = P(X \le 0) = P(X=0) = 1-\theta$

$F_X(0.5) = P(X \le 0.5) = P(X=0) = 1-\theta$

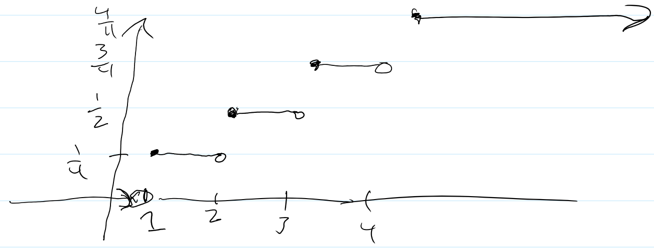
$F_X(0.7) = 1-\theta$

$F_X(1) = P(X \le 1)$
$= P(X=0) + P(X=1)$
$= 1-\theta + \theta$

$F_X(1.2) = P(X \le 1.2)$
$= 1$

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

Plot the graph of $F_X(x)$



$F_X(0.5) = P(X < 0.5)$
$= 0$

$F_X(1) = P(X \le 1)$

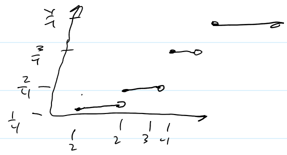
$F_X(2) = P(X \le 2) = P(X=2) + P(X=1)$
$= \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = \frac{1}{2}$

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

Rooms AA112, AC223

Last-hour: cdf's

Ex $P_X(x) = \frac{1}{4}$ $x \in \{1,2,3,4\}$
$= 0$ o/w



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

$F_X(0.5) = 0$
$F_X(2) = \frac{2}{4} = \frac{1}{2}$

$F_X(2^-) = \lim_{n\to\infty} F_X(2-\frac{1}{n}) = \frac{1}{4} \ne F_X(2)$

$F_X(2+) = \lim_{n\to\infty} F_X(2+\frac{1}{n}) = \frac{2}{4} = \frac{1}{2} = F_X(2)$

$f(x)$ is continous at a if $\lim_{x\to a} f(x) = f(a)$

$\lim_{x\to a^-} f(x) = f(a)$, left continous

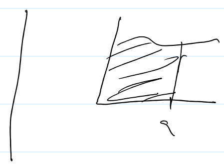
$\lim_{x\to a^+} f(x) = f(a)$, right continous

F_X is right continous at 2 but F_X is not left continous at 2

Note: The cdf's are right continous and have finite left limits. => [cadlag functions]

CDF's of continuous random variables

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)\,dt$$



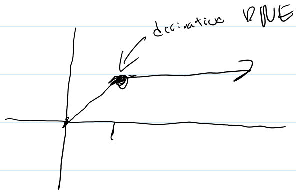
$$f_X(x) = \frac{d}{dx} F_X(x)$$

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

Find the pdf of X.

$1, \quad x > 1$

Find the pdf of x.

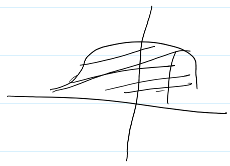


$$f_x(x) = \frac{d}{dx} F_x(x) = \begin{cases} 0 & , x < 0 \\ 1 & , 0 \le x \le 1 \\ 0 & , x > 1 \end{cases}$$ ← pay attention $<$ vs $\le$ be careful.

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

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

$$= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\,dt$$ ← only table calc can't help



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

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

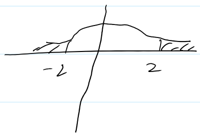
For any cont $< = \le$

$P(X \in (a,b]) = P(a < X \le b) = F_X(b) - F_X(a)$

$$= \Phi(2.00) - \underbrace{\Phi(-0.63)}_{=0.2643}$$

Table on pg 712

Z	0.0	.01	.0.03	... 0.09
−3.4	:		:	
−2.3	:		:	
−2.4	0.0228			
−0.6			0.2643	
:				
−0.0				



area z = area $-z$

$$\boxed{\Phi(x) = 1 - \Phi(-x)}$$

$$= \Phi(2.00) - \underbrace{\Phi(-0.63)}_{=0.2643}$$

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

$$= 1 - 0.0228 - 0.2643$$

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

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

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

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

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

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

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

$$= \left[\Phi(z) \right]_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}}$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

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)

$y \sim N(\underset{\mu}{-8}, \underset{\sigma^2}{4})$, Find (a) $P(\underset{a}{-2} < y \leq \underset{b}{7})$

(b) $P(y \geq 3)$

a) $P(-2 < y \leq 7) = \Phi\left(\frac{7-(-8)}{2}\right) - \Phi\left(\frac{-2-(-8)}{2}\right)$

$= \underset{0}{\Phi(7.5)} - \underset{0.0013}{\Phi(3)}$

$= 0.0013.$

	0.06 ... 0.07
-3.0	0.0013
0.0	

b) $P(y > 3) = 1 - P(y \leq 3) = 1 - \Phi\left(\frac{3-\mu}{\sigma}\right)$

$= \Phi(\infty) - \Phi\left(\frac{3--8}{2}\right)$

$= 1 - \Phi(5.5)$

$= \cancel{1} - (1 - \overset{0}{\Phi(-5.5)})$

$= 0$

$P(y > 3)$

$= P(\underset{a}{3} < y < \underset{b}{\infty})$

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

Rooms AA112, AC223

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 $F_Y(y) = P(Y \leq y) = P(h(X) \leq y)$ etc... ⟵ method has no conditions.

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

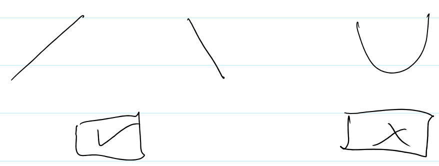
Today (Method 2: One dimensional transformation formula)

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

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

Thrm 2.6.3

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



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

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

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

Example: A random variable X has pdf

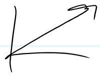
$f_X(x) = \begin{cases} 2(1-x) & , 0 \leq x \leq 1 \end{cases}$

$\hookrightarrow 0$ 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 $h^{-1}(y) = \frac{1+y}{2}$ ✓

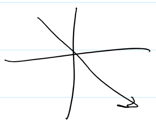
③ $h'(x) = \frac{dh(x)}{dx} = 2$

$$f_y(y) = \frac{f_x(h^{-1}(y))}{|h'(h^{-1}(y))|} = \frac{\cancel{2}(1 - \frac{1+y}{2})}{\cancel{|2|}}, \quad 0 \le \frac{1+y}{2} \le 1$$

$$= \left(\frac{1-y}{2}\right), \quad -1 \le y \le +1$$

$$= 0 \quad \text{o/w}$$

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

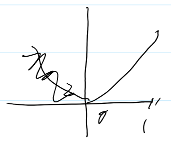
①  ✓ ② $h^{-1}(y) = \frac{1-y}{2}$ ③ $h'(x) = \frac{dh(x)}{dx} = -2$

$$f_y(y) = \frac{f_x(h^{-1}(y))}{|h'(h^{-1}(y))|} = \frac{2\left(1 - \frac{1-y}{2}\right)}{|-2|} \quad 0 \le \frac{1-y}{2} \le 1$$

$$= \left(1 - \frac{1-y}{2}\right), \quad 0 \le \frac{1-y}{2} \le 1$$

$$= 0 \quad \text{o/w}$$

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



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

$f_Y(y) = \frac{f_X(h(y))}{|h'(h^{-1}(y))|} = \frac{2(1-\sqrt{y})}{2\sqrt{y}}$, $0 \le \sqrt{y} \le 1$

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

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

$= 0$, o/w

2.7 joint distributions

Definition:

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

The joint distribution of two random variables X and Y is the collection of probability

$P((X,Y) \in B)$, $\forall B \subseteq \mathbb{R}^2$

$(X,Y) \in B = \{s : X(s), Y(s)) \in B\}$

Definition joint cdf:

The joint cdf of random variables X & Y is the function $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ defined by
$F_{X,Y}(x,y) = P(X \le x, Y \le y)$

recall,
cdf $F_X(x) = P(X \le x)$
$F_X : \mathbb{R} \to [0,1]$

Marginal distribution

$F_X(x) = \lim_{y \to \infty} F_{XY}(x,y) = F_{XY}(x,\infty)$

$F_Y(y) = \lim_{x \to \infty} F_{XY}(x,y)$ " $F_{XY}(\infty, y)$

"," = $\wedge$ = and

Notes
1. $F_{X,Y}(\infty,\infty) = 1$
2. $F_{X,Y}(-\infty,-\infty) = 0$
3. $F_{XY}(-\infty, y) = 0$
4. $F_{XY}(x, -\infty) \ge 0$

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 $P_{XY}: \mathbb{R}^2 \to [0,1]$ defined by $P_{XY}(x,y) = P(X=x, Y=y)$

$P_X(x) = e^{-\lambda}\frac{\lambda^x}{x!}$, $x = 0, 1, \ldots$
$= 0$ o/w

x	y
1	0.1
2	0.2
10	0.7

Consider an experiment of tossing a fair coin and a six sided die. X = outcome of the die, Y = # H's on the coin (0,1). Find the joint pmf of X and Y

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

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

July 3, 2018 10:04 AM

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

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

Marginal PDF of x: $P_X(x) = \sum_y P_{xy}(x,y)$ | $P_Y(y) = \sum_x P_{xy}(x,y)$

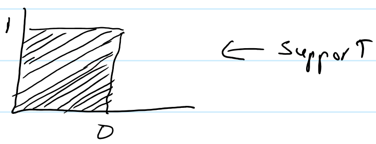
Joint density function satisfies

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

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

Ex (done)

$$f_{xy}(x,y) = \begin{cases} Kxy, & 0 \leq y \leq 1 \\ 0, & o/w \end{cases}$$



(a) Find K (Ans K = 4)

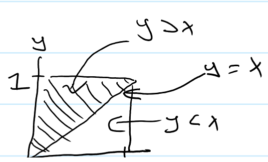
(b) Find $P(X \leq \frac{1}{2}, Y \leq \frac{3}{4})$

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

Today

Ex Random variables x + y have joint pdf

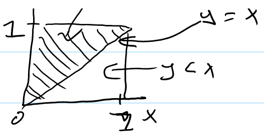
$$f_{x,y}(x,y) = \begin{cases} K(1-y), & 0 \leq x \leq y \leq 1 \\ 0, & O/W \end{cases}$$



Find K

$\infty \quad \infty$

Find K



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

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

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

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

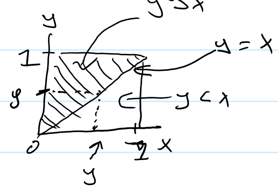
$$= K\int_0^1 (1-y)\,[x]_0^y\,dy$$

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

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

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

To figure out limit show support



Ex/ integrate the integral by changing dy dx

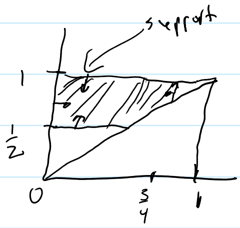
Part B

common = $\cap$ = intersection

Find the $P(X \le \frac{3}{4},\ Y \ge \frac{1}{2})$

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

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

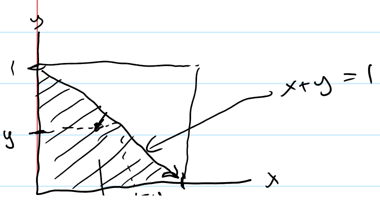


$$= 6\left[\frac{y^2}{2}-\frac{y^3}{3}\right]_{\frac{1}{2}}^{\frac{3}{4}} + 6\times\frac{3}{4}\left[y-\frac{y^2}{2}\right]_{\frac{3}{4}}^{1}$$

$= ?$ Ex

Ex $f_{xy}(x,y) = \begin{cases} 4xy, & 0 \le x \le 1,\ 0 < y \le 1 \\ 0, & o/w \end{cases}$

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

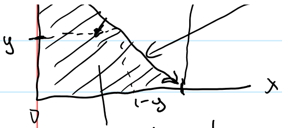


$$P(x+y \le 1) = \int_0^1\int_0^{1-y} 4xy\,dx\,dy$$

$$= 4\int_0^1 y\left[\frac{x^2}{2}\right]_0^{1-y} d$$

Recall

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



$x + y \le 1$

$x + y = 1$
then
$x = 1 - y$

$$= 4 \int_0^1 y \left[\frac{x^2}{2} \right]_0^{1-y} dy$$

$$= \frac{4}{2} \int_0^1 y(1-y^2)\, dy$$

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

$$= \frac{\cancel{2 \times 1 \times 2}}{4 \cdot 3 \cdot 2 \cdot 1} \times 2 =$$

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

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

Final exam type question

Ex Random variable x and y has joint density

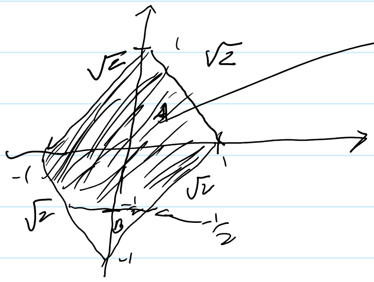
$$f_{xy}(x,y) = \begin{cases} k, & \text{if } |x| + |y| \le 1 \\ 0, & 0/w \end{cases}$$

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

you don't need to find k.

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

$|x| + |y| \le 1$



any point is equally likely to receive the point

Marginal distributions

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

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

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

Theorem 2.7.5 If x and y are jointly absolute continuous random variables with joint pdf f(x,y) then the marginal pdf's are given by

$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$, if $f_X(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$

Ex $f_{XY}(x,y) = \begin{cases} 4xy & , 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & , o/w \end{cases}$

Find f_X: ⟵ marginal pdf of f_X

Answer

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

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

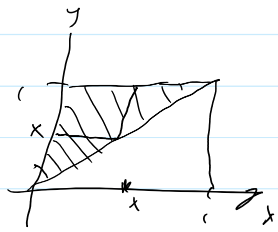
remember this

$= 0$, o/w

Ex $f_Y(y)$? (ans $2y$) , $0 \le y \le 1$

Ex $f_{XY}(x,y) = \begin{cases} 6(1-y) & , 0 \le x \le y \le 1 \\ 0 & , o/w \end{cases}$

Find f_X



Ans $f_X(x) = \int_x^1 6(1-y)\,dy$

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

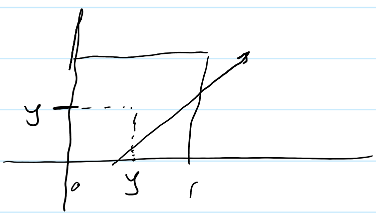
$= 3(1-x)^2$, $0 < x < 1$

$= 0$, o/w

b) Find $f_Y(y)$

$f_Y(y) = \int 6(1-y)\,dx$

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



$= 6(1-y)[x]_0^1$

$= 6y(1-y) \quad 0<y<1$

$= 0 \quad o/w$

2.8 Conditional distributions

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

$$\frac{P(a<Y\le b, X=x)}{P(X=x)}, \quad \text{for the event } a<Y\le b$$

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

$$P_{Y|X}(y|x) = \frac{P_{XY}(x,y)}{P_X(x)} = \frac{P(X=x, Y=y)}{P(X=x)}$$

Example

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

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

The joint mass function of X and Y

$x \backslash y$	0	1	
0	0.38	0.17	0.55
1	0.14	0.02	0.16
2	0.24	0.05	0.29
	.76	.24	

a) Give the marginal pmf's of X and Y

x	$p_X(x)$
0	0.55
1	0.16
2	0.29

$= 1$

$\Leftarrow$ marginal pmf of X

y	$p_Y(y)$
0	.76
1	.24

$\Leftarrow$ marginal pmf for Y

b) Give the condition p' mass function for Y given $X=0$.

Ans $p_{Y|X}(y|0) = \dfrac{p_{XY}(0,y)}{p_X(0)}$

y	$p_{Y\|X}(y\|0)$
0	$\frac{0.38}{0.55}$
1	$\frac{0.17}{0.55}$

$\dfrac{p_{XY}(0,1)}{p_X(0)}$

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 $f_{Y|X}(y|x) = f_{XY}($

July 10, 2018 10:04 AM

Last hour: Section 2.8 conditional distribution

Definition if x and y are jointly discrete

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

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 absolute continuous

The conditional pdf: $f_{Y|X}(y|x) = \dfrac{f_{XY}(x,y)}{f_X(x)}$, $f_X(x)>0$

Definition $P(a\le Y\le b \mid X=x) = \int_a^b f_{Y|X}(y|x)\,dy$

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 $P(a\le X\le b, c\le Y\le d) = P(a\le X\le b)\cdot P(c\le Y\le d)$
whenever $a\le b$, $c\le d$.

Thrm 2.8.3

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

$$P_{xy}(x,y) = P_x(x) \cdot P_y(y) \quad \forall x,y \in \mathbb{R}$$

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

$$f_{xy}(x,y) = f_x(x)\, f_y(y) \quad \forall x,y \in \mathbb{R}$$ $\longleftarrow$ This ones easiest to prove $\perp$

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

Are x and y $\perp$?

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

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

$$f_{xy}(x,y) = 4xy = (2x)(2y) = f_x(x)\, f_y(y)$$

$\therefore\ x \perp y \quad \blacksquare$

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

Are x and y $\perp$?

$f_x(x) = 3(1-x)^2,\ 0 \leq x \leq 1$

$f_y(y) = 6y(1-y),\ 0 \leq y \leq 1$

$x = \frac{1}{2},\ y = \frac{1}{2}$

$f_X(\frac{1}{2}, \frac{1}{2}) = 6(1-\frac{1}{2}) = 3$

$\cdot\ f_X(\frac{1}{2}) \times f_y(\frac{1}{2}) = 3[1-\frac{1}{2}]^2 \times 6 \times \frac{1}{2}(1-\frac{1}{2})$

$\neq 3$

$\therefore X \not\perp Y$ ← not independent

Thrm 2.8.4

a) if X and Y are jointly discrete then
$X \perp Y$ iif $P_{Y|X}(y|x) = P_Y(y) \quad \forall x, y \in \mathbb{R}$ s.t $P_X(x) > 0$

b) if X and Y are jointly absolutely continuous, then
$X \perp Y$ iif $f_{Y|X}(y|x) = f_Y(y) \quad \forall x, y \in \mathbb{R}$ 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

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

Ex/ A r.v X has pmf

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

Find $E(X)$

Ans $E(X) = \sum_x x p_X(x) = (1 \times 0.2) + (2 \times 0.3) + (3 \times 0.5)$

① $X \sim$ degenerate at c, find $E(X)$?

Recall Degenerate distribution:

x	$P_X(x)$
c	1

mean of $X = E(X) = c \times 1 = c$

② $X \sim Ber(\theta)$. Find $E(X)$

Ans:

x	$P_X(\theta)$
0	$1-\theta$
1	θ

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

③ $X \sim Bin(n, \theta)$. Find $E(X)$

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

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

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

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

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

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

let $y = x - 1$

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

recall $(a+b)^n$

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

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

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

$= n\theta$

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

<u>Basic Math</u> (Recall)

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

radius of convergence $\mathbb{R}$, 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.

$$f'(x) = \sum_{n=1}^{\infty} a_n \cdot n(x-c)^{n-1}$$

$$f''(x) = \sum_{n=2}^{\infty} a_n n(n-1)(x-c)^{n-2}$$

$$\frac{d}{dx} \sum a_n (x-c)^n$$

$$= \sum \frac{d}{dx} a_n (x-c)^n$$

etc...

(*) $x \sim Geo(\theta)$, Find $E(x)$? $\quad p_x(x) = (1-\theta)^x \theta, \; x = 0, 1, 2, \ldots$

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

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

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

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

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

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

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

$$= \frac{a}{1-r}$$

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

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

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

$$= \theta q \left(0 + (1-q)^{2}\right]$$

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

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

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

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

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

Expect refers 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 = # of tails before the first head

then

$X \sim Geo(\theta = 0.1)$

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

(5) $X \sim Po(\lambda)$. Find $E(X)$: Recall $p_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ $x = 0, 1, 2, \ldots$

Ans $E(X) \overset{def}{=} \sum_x x\, p_X(x) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$, $e^{-\lambda}$ is constant

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

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

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

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

$$= \lambda e^{-\lambda} \left[\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{3!} + \dots \right]$$

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

$$= \lambda$$

$X \sim Po(\lambda) \Rightarrow E(X) = \lambda$

remember λ is mean.

Ex A random variable X has pmf $p_X(x) = \frac{1}{x(x+1)}$, $x = 1, 2, 3, \dots$

Find $E(X)$

Ans $E(X) = \sum_x x \, p_X(x) = \sum_{x=1}^{\infty} x \, p_X(x)$

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

$$= \infty$$

expected to know

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

Ex show that a pmf i.e

① $p_X(x) \geq 0, \forall x$

② $\sum_{n=1}^{\infty} \frac{1}{x(x+1)} = 1$

Ex A r.v X has pmf

$$p_X(x) = \begin{cases} \frac{1}{2x} & x = 2, 4, 8, 16 \quad (2^1, 2^2, 2^3, 2^4 \dots 2^n) \\ \frac{1}{2|x|} & x = -2, -4, -8, -16 \dots -2^n \end{cases}$$

Find E(x)

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

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

$$= \infty - \infty$$

$$= \infty$$

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

Last hour   Chapter 3 Expectation/mean/expected value

- Discrete random variables [1]
- Definition $E(X) = \sum_x P_X(x)$
- $X \sim Degenerate(c) \Rightarrow E(X) = c$
- $X \sim Ber(\theta) \Rightarrow E(X) = \theta$
- $X \sim Bin(n, \theta) \Rightarrow E(X) = n\theta$
- $X \sim Geo(\theta) \Rightarrow E(X) = \frac{1-\theta}{\theta}$
- $X \sim P_o(\lambda) \Rightarrow E(X) = \lambda$
- $X \sim f_X(x) = \frac{1}{x(x+1)}$, $x = 1, 2, \dots \Rightarrow E(X) = \infty$
- $X \sim f_X(x) = \begin{cases} \frac{1}{2x}, & x = 2, 4, 8, \dots 2^n \\ \frac{1}{2(x)}, & x = -2, -4, -8, \dots -2^n \end{cases} \Rightarrow$ $E(X)$ does not exist if $E(X) = \infty - \infty$

$y = h(x)$

f_X

$f_Y = \sum_y y P_X(x)$

E_y

Thrm 3.1

(a) Let X be a discrete random variable and $h: \mathbb{R} \to \mathbb{R}$ 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 $h: \mathbb{R}^2 \to \mathbb{R}$ 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	$P_X(x)$
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) $E\left[\frac{1}{X}\right] = \sum_x \frac{1}{x} P_X(x) = \frac{1}{1} \cdot 0.5 + \frac{1}{2} \times 0.5 = 0.75$

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

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

$E(x)$ $\frac{\ }{2}$

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

x \ 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 $E(y)$

Ans a) $E(xy) = 0 \times 0 \times .2 + 0 \times 1 \times 0.1 + 1 \times 0 \times 0.4 + 1 \times 1 \times 0.3 = 0.3$

b) $E(x) = \sum x P_x(x) = 0 \times 3 + 1 \times 0.7 = 0.7$

c) $E(y) = \sum_y y P_y(y) = 0 \times 0.6 + 0.4 = 0.4$

Note: $E(x) \times E(y) != E(xy)$

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?

Ans $x \sim Geo(\theta = \frac{1}{2}) \Rightarrow P_x(x) = (1-\theta)^x \theta$, $x = 0, 1, 2$
$= (\frac{1}{2})^{x+1}$, $x = 0, 1, \ldots$

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

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

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

b) $P(2^x \geq \$256)$
$= P(2^x \geq 2^8)$
$= P(x \geq 8)$
$= P_x(8) + P_x(9) + P_x(10) + \ldots$

$P_x(x) = \left(\frac{1}{2}\right)^{x+1}$ $x = 0, 1, 2, \ldots$

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

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

Thm 3.1.2 Linearity of Expectation

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

pf $E[ax+by] = \sum_x \sum_y (ax+by) P_{xy}(x,y)$

$= \sum_x \sum_y (ax P_{xy}(x,y) + by P_x(x,y))$

$= \sum_x ax \sum_y P_{xy}(x,y) + \sum_y by \sum_x P_{x,y}(x,y)$

$= a \sum_x x P_x(x) + b \sum_y y P_y(y)$

$= aE(x) + bE(y)$, by definition ▨

This also implies,

① $E(x+y) = E(x) + E(y)$, but product NO!

② $E[ax] = aE(x)$

▨

July 17, 2018 10:04 AM

Last hour: (3.1 Expectation)

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

- $x \sim \text{Degenerate}(c) \Rightarrow E(x) = c$
- $x \sim \text{Ber}\,\theta \Rightarrow E_x = \theta$
- $x \sim \text{Bin}(n, \theta) \Rightarrow E_x = n\theta$
- $x \sim P_o(\lambda) \Rightarrow E_x = \lambda$
- $x \sim \text{Geo}\,\theta \Rightarrow E_x = \frac{1-\theta}{\theta}$

- $E(x)$ can be ∞ ($\pm\infty$)
- $E(x)$ can be non-existent (i.e $\infty - \infty$)
- $E(h(x)) = \sum_x h(x) p_x(x) \ldots$ i.e we do not need the pmf of $h(x)$ to calculate $E(h(x))$
- E is a linear operator

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

$\Rightarrow E(ax) = aE(x) \ldots$ (taking $b=0$)

$\Rightarrow E(x+b) = E(x) + b$ taking ($a=1$ and $y=1$)

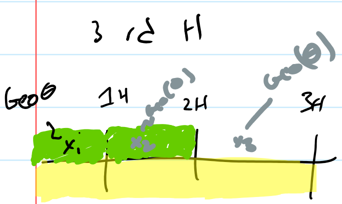
$\Rightarrow E(x+y) = E(x) + E(y)$

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

Note

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

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.

$\therefore$ when y is negative binomial:

iid $\Rightarrow$ independent and identically distributed

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

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

Recall

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

Thrm if x and y are independent then

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

<u>Proof (Discrete case)</u>

$$E[g(x)h(y)] = \sum_x \sum_y g(x)h(y)\,p_{x,y}(x,y)$$

<u>Recall</u>

$$E(g(x,y)) = \sum_x \sum_y g(x,y)\,P_{xy}(x,y)$$

$$x \perp y \Rightarrow P_{xy}(x,y) = P_x(x)P_y(y)$$

$$= \sum_x \sum_y g(x)h(y)p_x(x)\cdot p_y(y)$$

$$= \sum_x g(x)\,p_x(x) \sum_y h(y)p_y(y)$$

$$= \underbrace{\sum_y h(y)p_y(y)}_{E(h(y))} \underbrace{\sum_x g(x)\,p_x(x)}_{E(g(x))}$$

$$= E(g(x))\cdot E(h(y)) \quad \blacksquare$$

<u>Note</u> if $x \perp y \Rightarrow E(xy) = E(x)E(y)$ (!) only if independent

$$E(x+y) = E(x) + E(y)$$ always true ✓

<u>Thrm 3.1.4</u> Monotonicity of expectations

if $x \le y \Rightarrow E(x) \le E(y)$

$x \le y \iff x(s) \le y(s)$ $\forall s \in S$

<u>Proof</u>

Define: $Z = y - x$ then
$$Z(s) = Y(s) - X(s) \ge 0 \quad \forall s \in S$$

Define: $Z = y - x$ then

$$Z(s) = Y(s) - X(s) \geq 0 \quad \forall s \in S$$

$$E(y-x) = E(Z) = \sum_{z} \underset{\substack{| \\ z \geq 0}}{z}\; P(Z=z)$$

$P(Z=z) \geq 0$

$P_Z(z) = P(Z=z)$

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

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 absoulutely continuous random variable then its expectations is given by:

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

Discrete case

$$E_X = \sum_{x} x\, p_X(x)$$

$$= \int x f_X(x)\, dx$$

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

Recall $f_X(x)$ of uniform

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

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

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

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

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

$$= \frac{1}{R-L} \left[\frac{x^2}{2} \right]_L^R$$

$$= \frac{1}{R-L} \left(\frac{R^2-L^2}{2} \right) = \frac{L+R}{2} \qquad \blacksquare$$

<u>Ex2.</u> $X \sim \exp(\lambda)$. Find $E(X)$?

<u>Ans</u>

$$E(X) = \int_0^\infty x f_X(x)\,dx = \int_0^\infty \underbrace{x}_{u} \underbrace{\lambda e^{-\lambda x} dx}_{dv}$$

<u>Recall</u>

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$
$$= 0, \quad \text{o/w}$$

$u = x \qquad du = dx$

$v = -e^{-\lambda x} \qquad dv = \lambda e^{-\lambda x}$

$$\left[-x e^{-\lambda x}\right]_0^\infty + \int_0^\infty e^{-\lambda x}\, dx$$

$$= 0 + \frac{1}{\lambda} = \frac{1}{\lambda}$$

$$\lim_{x \to \infty} \frac{x}{e^{\lambda x}} = 0$$

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

<u>Example 3</u> $Z \sim N(\underset{\mu}{0}, \underset{\sigma^2}{1})$. Find $E(Z)$

<u>Recall</u>

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad -\infty < z < +\infty$$

<u>Ans</u>:

$$E(Z) = \int_{-\infty}^{\infty} z\, f_Z(z)\, dz$$

$$= \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz$$

$$= \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^{0} z\, e^{-\frac{1}{2}z^2}\, dz + \int_0^\infty z\, e^{-\frac{1}{2}z^2}\, dz \right]$$

$$= \frac{1}{\sqrt{2\pi}} \left[\left[-e^{-\frac{1}{2}z^2}\right]_{-\infty}^{0} + \left[-e^{-\frac{1}{2}z^2}\right]_0^\infty \right]$$

$= \frac{1}{\sqrt{2\pi}} \left\{ \left[-e^{-\frac{z}{2}} \right]_{-\infty}^{} + \left[-e^{z\,0} \right]_{0} \right\}$

$= \frac{1}{\sqrt{2\pi}} \left[(-1 - (0)) + (0 - (-1)) \right]$

$= 0$

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

Thrm 3.2.1

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

$$E(g(x)) = \int g(x) \cdot f_x(x)\, dx$$

$y = g(x)$

$E_y = \int y\, f_y(y)\, dy$

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

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

$y = h(x,y)$

$E y = \int y\, f_y(y)\, dy$

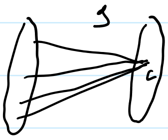
Thrm 3.2.2 (Linearity of expectation)

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

$\Rightarrow E[ax] = aE(x)$

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

Ex x is a absolutely continuous random variable with pdf f_x



$g(x) = c$

Find $E(g(x))$?

Soln

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

$$= \int_{-\infty}^{\infty} c \; f_X(x)\,dx$$

$$= c \underbrace{\int_{-\infty}^{\infty} f_X(x)\,dx}_{1}$$

1 because it's a density function and by def $\int_{-\infty} f_X(x)\,dx = 1$

$$= c$$

$\therefore \; E(c) = c$

Ex $X \sim N(\mu, \sigma^2)$. Find $E(X)$

Soln $X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

$$0 = E(Z) = E\left[\frac{X-\mu}{\sigma}\right]$$

$$- \quad \sigma \quad - \frac{\mu}{\sigma}]$$

$$= \frac{1}{\sigma} E(X) - \frac{\mu}{\sigma} \Rightarrow E(X) = \mu$$

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

[Ans: $\frac{X-\mu}{\sigma} \sim N(0,1)$

distribution

$\left[\text{Ans: } \frac{x-\mu}{\sigma} \sim N(0,1) \right.$

Ex $X \sim f_X(x) = \frac{1}{x^2}$, $x > 1$

$= 0$, o/w

show its a pdf

$\frac{(+)}{(+)}$ and $\int_{-\infty}^{\infty} f_X(x) = 1$

Find $E(x)$?

Soln $E(x) = \int_1^{\infty} x f_X(x) dx$

$= \int_1^{\infty} x \frac{1}{x^2} = \int_1^{\infty} \frac{1}{x} dx = [\ln|x|]_1^{\infty}$

$= \infty - 0 = \infty$

Ex $X \sim f_X(x) = \begin{cases} \frac{1}{2x^2} & , x > 1 \\ \frac{1}{2x^2} & , x < -1 \\ 0 & , o/w \end{cases}$

Find $E(x) = \int_{-\infty}^{-1} x f_X(x) dx + \int_1^{\infty} x f_X(dx)$

$= \frac{1}{2} \int_{-\infty}^{1} x \frac{1}{x^2} dx + \frac{1}{2} \int_1^{\infty} x \frac{1}{x^2} dx$

$= \frac{1}{2} [\ln|x|]_{-\infty}^{1} + \frac{1}{2} [\ln|x|]_1^{\infty}$

$= \frac{1}{2} [0 - \ln|-\infty|] + \frac{1}{2} [\ln|\infty| - 0]$

$= \frac{1}{2} [0 - \infty] + \frac{1}{2} [\infty - 0]$

$= -\infty + \infty = \infty - \infty \Rightarrow DNE$

Thrm 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 $Var(x) = E[(x-\mu_x)^2]$

Note: $Var(x) \geq 0$

Result: $Var(x) = E[x^2] - [E(x)]^2$ or $(E(x^2) - \mu_x^2)$

PF $Var(x) = E[(x-\mu_x)^2]$

$E(ax) = aEx$

$$= E[x^2 - 2\mu_x x + \mu_x^2]$$

$$= E(x^2) - 2\mu_x \underbrace{E(x)}_{\mu_x} + \mu_x^2$$

$$= E(x^2) - 2\mu_x^2 + \mu_x^2$$

$$= E(x^2) - \mu_x^2 = E(x^2) - (E(x))^2$$

Note $(E(x))^2 \leq E(x^2)$ because $var(x) \geq 0$

July 12, 2018 11:13 AM

Last hour: Section 3.3 variance

Definition: $Var(x) = E[(x-E(x))^2]$ ← harder

$E(x)$
$E(x^2)$

Results $Var(x) = E(x^2) - (E(x))^2$ ← easier

Notes: $Var(x) \geq 0 \quad \forall$ r.v x

$(E(x))^2 \leq E(x^2), \quad \forall$ r.v x

Before var

- $x \perp y \Rightarrow E[g(x)h(y)] = E(g(x)) \cdot E(h(y))$

$(x \perp x \Rightarrow E(xy) = Ex \cdot Ey)$

- $x \leq y \Rightarrow E(x) \leq E(y)$ monotonicity of expectation
- $E(ax+by) = aE(x) + bE(y)$ linearity

Definition: Standard deviation of a r.v x

$SD(x) = +\sqrt{Var(x)}$, (Denoted by σ_x)

Results

(1) $Var(c) = 0$, since a constant doesn't vary. It is fixed.

pf: $Var(c) = E[(c-E(c))^2]$

$= E[(c-c)^2]$
$= E[0] = 0$

(2) $Var(a+bx) = b^2 Var(x)$

Pf $Var(a+bx) = E[(a+bx) - E(a+bx)^2]$

$$= E[((a+bx - a - bE(x))^2]$$

$$= E[(b(x-E(x))^2]$$

$$= b^2 E[(x-E(x))^2]$$

$$= b\,Var(x)$$

This result implies:

- $Var(a+x) = Var(x)$
- $Var(bx) = b^2 var(x)$ $\Leftarrow$ remember of variance when you take a constant out, to square it.
- $SD(a+bx) = +\sqrt{var(a+bx)}$

$$= +\sqrt{b^2 var(x)}$$

$$= |b| SD(x)$$

$$= |b| + \sqrt{var(x)}$$

Ex $x \sim Ber(\theta)$, Find $var(x)$

Ans

$$Var(x) = E(x^2) - (E(x))^2$$

$$= \theta - (\theta)^2$$

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

$\therefore$ $x \sim Ber(\theta)$

$\Rightarrow Var(x) = \theta(1-\theta)$

$E(x) = \theta$

Recall

$E(x) = \theta$ for bernoulli

$p_x(x)$

x	p_x
0	$1-\theta$
1	θ

$E(x^2)$

$$= \sum_x x^2 p_x(x)$$

$$= 0^2 \cdot (1-\theta) + 1^2 \cdot \theta$$

$$= \theta$$

Ex $X \sim Bin(n, \theta)$. Find varience

Ans $Var(x) = E(x^2) - (E(x))^2$

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

you can change this to two

$$= \sum_{x=2}^{n} x(x-1)\frac{n!}{x!(n-x)}\theta^x(1-\theta)^{n-x}$$

$$= \sum_{x=2}^{n} (x-1)\frac{n!}{(x-2)(n-x)}\theta^x(1-\theta)^{n-x}$$

$$= n(n-1)\theta^2 \sum_{x=2}^{n} \frac{(n-2)!}{(n-2)!\,[(n-2)-(x-2)]!}\theta^{n-2}(1-\theta)^{(n-2)-(x-2)}$$

$$= n(n-1)\theta^2 \sum_{y=0}^{n-2} \frac{(n-2)!}{y!((n-2)-y)!}\theta^y(1-\theta)^{(n-2)-y}$$

$\binom{n-2}{y}\theta^y(1-\theta)^{(n-2)-y} \leftarrow$ binomial

$$= (\theta+1-\theta)^{n-2}\cdot n(n-1)\theta^2$$
$$= n^2\theta^2 - n\theta^2$$

$$E[x(x-1)] = n^2\theta^2 - n\theta^2$$

Recall

if $x \sim bin$ then
$E(x) = n\theta$

$\sum g(x)P_x(x)$

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

Sub $y = x-2$

$(a+b)^n = \sum \binom{n}{x} a^x b^{n-x}$

$a = \theta$

$b = 1-\theta$

$$E(x^2) - E(x) = n^2\theta^2 - n\theta^2$$

$$E(x^2) = n^2\theta^2 - n\theta^2 + n\theta$$

$$Var(x) = E(x^2) - (E(x))^2$$
$$= \cancel{n^2\theta^2} - n\theta^2 + n\theta - \cancel{(n\theta)^2}$$
$$= n\theta(1-\theta)$$

$\blacksquare$

recall $E(x) = n\theta$ for binomial

$x \sim Bin(n,\theta)$

$\Rightarrow Var(x) = n\theta(1-\theta)$

$\Rightarrow E(x) = n\theta$

Ex $X \sim P_o(\lambda)$, Find $Var(x)$

Ex $X \sim P_o(\lambda)$, Find $Var(x)$

Ans $Var(x) = E(x^2) - (E(x))^2$

$Ex^2 = \sum_{x=0}^{\infty} x^2 p_x(x) \leftarrow$ too messy
use factorial shortcuts

$E[x(x-1)] = \sum_{x=\cancel{0}\ 2}^{\infty} \cancel{x}\cancel{(x-1)} e^{-\lambda} \frac{\lambda^x}{\cancel{x!}\ (x-2)!}$

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

$= \cancel{e^{-\lambda}} \lambda^2 \cancel{e^{\lambda}}$

$= \lambda^2$

Recall

If $x \sim P_o$

$E(x) = \lambda$

$p_x(x) = e^{-\lambda} \frac{\lambda^x}{x!}$

change summation limits

use factorial shortcut

convert all to $(z - x)$ form

sub $y = (x-2)$ (not in this case though)

$E(x(x-1)) = \lambda^2$

$E(x^2) - E(x) = \lambda^2$

$E(x^2) - \lambda = \lambda^2$

$E(x^2) = \lambda^2 + \lambda$
$= \lambda(\lambda + 1)$

$Var(x) = E(x^2) - (E(x))^2$
$= \cancel{\lambda^2} + \lambda - \cancel{\lambda^2}$
$= \lambda$

$\therefore$ if $x \sim P_o(x) = Var(x) = \lambda$

$E(x) = \lambda$

if mean and variance = use poisson.

Last hour: Sec 3.3 Varience

* Definition: $Var(x) = E[(x-\mu_x)]$ where $\mu_x = E(x)$ ← Difficult to calculate Variance with the definition, use the result.
* Result: $Var(x) = E(x^2) - (E(x))^2$
* Standard deviation, $SD(x) = +\sqrt{var(x)}$
* Properties
  - $Var(x) \geq 0$
  - $Var(a-bx) = b^2 Var(x)$ ← remember to square the b^2
  - $Var(c) = 0 \quad \forall c \in \mathbb{R}$

Examples

$x \sim Ber(\theta) \Rightarrow Var(x) = \theta(1-\theta)$, $E(x) = \theta$

$x \sim Bin(n,\theta) \Rightarrow Var(x) = n\theta(1-\theta)$, $E(x) = n\theta$

$x \sim P_o(\lambda) \Rightarrow Var(x) = \lambda$, $E(x) = \lambda$

Today

Ex/ $x \sim Geo(\theta)$. Find $Var(x)$?

Ans $V(x) = E(x^2) - (E(x))^2$

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

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

$$= \theta \sum_{x=2}^{\infty} (x)(x-1)(q)^x \quad , \text{ let } q = 1-\theta$$

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

Hint

Ex^2 is difficult use $E(x(x-1))$

helps when (!) ← factorials are involved.

Recall

if $x \sim Geo(\theta)$ then

$E(x) = \frac{1-\theta}{\theta}$ $\quad p_x(x) = (1-\theta)^x\theta \quad , x = 0,1,2,\ldots$

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

$$= \theta q^2 \sum_{x=2}^{\infty} \frac{d^2}{dq} q^x$$

$$= \theta q^2 \frac{d^2}{dq} \sum_{x=2}^{\infty} q^x \leftarrow \text{geometric series } \frac{a}{1-r}$$

$$= \theta q^2 \frac{d^2}{dq}\left(\frac{q^2 - 1 + 1}{1-q}\right)$$

$$= \theta q^2 \frac{d^2}{dq}\left[-1 - q + \frac{1}{1-q}\right]$$

Recall

$\sum_{x=2}^{\infty} x(x-1)q^{x-2}$ ← power series

Interchange summation and derivatives

$$= \theta q^2 \left[\frac{2}{(1-q)^3} \right]$$

recall, $q = 1-\theta$

$\iff 1-q = \theta$

$$= \cancel{\theta} q^2 \left[\frac{2}{\theta^{\cancel{3}\,2}} \right] = \frac{2q^2}{\theta^2}$$

$$= \frac{2(1-\theta)^2}{\theta^2}$$

$$E(x^2) = \frac{2(1-\theta)^2}{\theta^2} + \frac{1-\theta}{\theta}$$

$$V(x) = E(x^2) - (E(x))^2$$

$$= \frac{\cancel{2}(1-\theta)^2}{\theta^2} + \frac{1-\theta}{\theta} - \cancel{\frac{(1-\theta)^2}{\theta^2}}$$

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

$$= \frac{1-\theta}{\theta^2}$$

$\therefore$ if $x \sim Geo(\theta)$ then

$\Rightarrow V(x) = \dfrac{1-\theta}{\theta^2}$ $\qquad E(x) = \dfrac{1-\theta}{\theta}$

Ex $x \sim Exp(\lambda)$ show that $Var(x) = \dfrac{1}{\lambda^2}$

Hints $Var(x) = E(x^2) - (E(x))^2$

$E(x^2) = \int_0^\infty x^2 \lambda^{-\lambda x} dx = \frac{2}{\lambda^2}$, show this

Recall

$E(x) = \dfrac{1}{\lambda}$

$f_X(x) = \lambda e^{-\lambda x}$ $\quad x \geq 0$

$u = \lambda x \qquad du = x\, dx$

$$= \int_0^\infty \frac{u}{\lambda} \lambda^{-u} du$$

$$= \int_0^\infty \frac{u}{\lambda \lambda^u} du$$

(homework excercise)

$$= \frac{1}{\lambda} \int_0^\infty \frac{u}{\lambda^u} du$$

Ex/ $z \sim N(0,1)$. Find $Var(z)$

$V(Z) = E(Z^2)$

$= \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$

$= \int_{-\infty}^{\infty} \underbrace{z}_{u} \cdot \overbrace{z \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz}^{dv}$

$E(z) = 0$

$Var(z) = E(z^2) - (E(z))^2$

$\Rightarrow Var(z) = E(z^2) - 0 = E(z^2)$

Let $u = z$, $dv = \frac{1}{\sqrt{2\pi}} z \cdot e^{-\frac{1}{2}z^2} dz$

$du = 1$, $v = \frac{1}{\sqrt{2\pi}}\left(-e^{-\frac{1}{2}z^2}\right)$

$= \frac{1}{\sqrt{2\pi}} \left[-z^2 e^{-\frac{1}{2}z^2}\right]_{-\infty}^{\infty} + \int_0^{\infty} e^{-\frac{1}{2}z^2} dz$

$= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = 1 \leftarrow$ since this is a density function the integral is = to 1

$Z \sim N(\underset{\mu}{0}, \underset{\sigma^2}{1}) \Rightarrow Var(Z) = 1,\ E(Z) = 0$

Recall $N(0,1)$ is standard normal and $N(\mu, \sigma^2)$ is normal.

Ex $x \sim N(\mu, \sigma^2) \Rightarrow z = \frac{x-\mu}{\sigma} \sim N(0,1)$

$z = \frac{1}{\sigma} x - \frac{\mu}{x}$

resembles $ax + b$

Recall

$V(ax+b) = a^2 v(x)$

$1 = V(Z) = \left(\frac{1}{\sigma^2}\right) V(x)$

$\Rightarrow V(x) = \sigma^2$

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

$E(x) = \mu$, $v(x) = \sigma^2$

Ex $x \sim Gamma(\alpha, \lambda)$. Find $var(x)$?

Ex $X \sim \text{Gamma}(\alpha, \lambda)$. Find $\text{var}(X)$?

Ans $V(X) = E(X^2) - (E(X))^2$

$$E(X) = \int_0^\infty \frac{x \lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} dx$$

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

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

Recall

Density

$$f_X(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x > 0$$

because its a density function

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

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

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

$$= \frac{\alpha}{\lambda}$$

$$\therefore E(X^2) = \frac{\alpha(\alpha+1)}{\lambda^2}$$

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}$$

Ex $X \sim \text{Beta}(a, b)$. Show that

(1). $E(X) = \frac{a}{a+b}$

(2). $\text{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}$

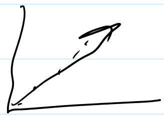
Last hour (Covariance, Correlation)

Definition

$$Cov(x,y) = E[(x-\mu_x)(y-\mu_y)]$$
$$= E(xy) - E(x)E(y)$$ ← easy way to calculate

Properties

- $Cov(x,y) = Cov(y,x)$
- $Cov(x, a+by) = b\,Cov(x,y)$ — prove this, exercise
- $Cov(x_1+x_2, y) = Cov(x_1,y) + Cov(x_2,y)$
- $Cov(x,x) = V(x)$
- $Var\left(\sum_{i=1}^{\infty} x_i\right) = \sum_{i=1}^{k} Var(x_i) + 2\sum_{i=j}\sum Cov(x_i,x_j)$
- $Corr(x,y) = \dfrac{Cov(x,y)}{\sqrt{Var(x)Var(y)}}$



- if $y = a+bx$ then $corr(x,y) = \dfrac{b}{|b|}$

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

① $r_x(t) = E(t^x) \Leftarrow$ definition

probability generating function

recall

$t^x = e^{\ln t}$

$r_x(t) = E(t^x) = E(e^{\ln(t)^x})$

$= E[e^{\ln t \, x}] = m_x(\ln t)$

② $m_x(t) = E[e^{tx}] = E[(e^t)^x] = r_x(e^t)$

Ex $x \sim Bin(n, \theta)$. Find $r_x(t)$ and $M_x(t)$

$r_x(t) = E[t^x] = \sum t^x p_x(x)$

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

recall

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

$x = 0, \dots n$

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

binomial

$(a+b)^n = \sum \binom{n}{x} a^x b^{n-x}$

$= (1-\theta + \theta t)^n$

$\therefore$ for $bin(n, \theta)$

$r_x(t) = (1-\theta + \theta t)^n \quad \forall t \in \mathbb{R}$

$m_x(t) = (1 - \theta + \theta e^t)^n$

Exercise, find it using first principle.

Ex $m_x(t) = E(e^{tx}) = (1-\theta + \theta e^t)^n$ show this.

Ex $x \sim Po(\lambda)$, show that $r_x(t) = e^{\lambda(t-1)}$, $t \in \mathbb{R}$

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

Ex $X \sim Geo(\theta)$, show that $\delta_X(t) = \dfrac{\theta}{1-(1-\theta)t}$

$$m_X(t) = \frac{\theta}{1-(1-\theta)(e^t)}$$

Ex $X \sim Exp(\lambda)$

since its a cont we aren't as interested in prob generating but rather moment generating function

$$m_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} f_X(x)\, dx$$

$$= \int_0^\infty e^{tx} e^{-\lambda x}\, dx$$

$$= \lambda \int_0^\infty e^{-(\lambda - t)x}\, dx$$

$$= \frac{\lambda}{\lambda - t}, \quad t < \lambda$$

Ex $Z \sim N(0,1)$ Find moment generating function of Z

$$m_Z(t) = E(e^{tZ}) = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2 + tz + t^2 - t^2}\, dz$$

$$= e^{\frac{1}{2}t^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2}\, dz}_{\text{normal distribution with mean } t. \text{ Integrate density} = 1}$$

this correct above couldn't see.

$$= e^{\frac{1}{2}t^2}$$

$\therefore$ if $Z \sim N(0,1) \Rightarrow m_Z(t) = e^{\frac{1}{2}t^2}$, $t \in \mathbb{R}$

<u>Result</u> if $y = a + bx$ then $m_y(t) = e^{at} m_x(bt)$

<u>Pf</u> $m_y(t) = E[e^{ty}] = E[e^{t(a+bx)}] = E[e^{at + btx}]$

$= \underbrace{e^{at}}_{\text{constant}} E[e^{btx}]$

$= e^{at} m_x(bt)$ $\blacksquare$

$m_x(t) = E(e^{tx})$

<u>Ex</u> $x \sim N(\mu, \sigma^2)$. Find m_x?

we know if $x \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{x - \mu}{\sigma^2} \sim N(0,1)$

$\Rightarrow$ x has the same distributio
$x = \mu + \sigma Z$
where Z is standard
normal ($Z \sim N(0,1)$)

$m_x(t) = e^{\mu t} m_Z(\sigma t)$

$= e^{\mu t} \cdot e^{\frac{1}{2}\sigma^2 t^2}$

$= e^{\mu t + \frac{1}{2}\sigma^2 t^2}$ $\blacksquare$

$m_Z(t) = e^{\frac{1}{2}t^2}$

<u>Results</u>

$r_x(t) = E[t^x]$

if x takes values $0, 1, 2 \ldots$

$E\, g(x)$
$= \sum_x g(x) p_x(x)$

$= g(0) p_x(0) + g(1) p_x(1) \ldots$

$r_X(t) = E[t^X]$

$E\, g(X)$
$= \sum_x g(x) p_X(x)$
$= g(0)p_X(0) + g(1)p_X(1) \dots$

if X takes values $0, 1, 2 \dots$

then

$r_X(t) = E[t^X]$

$= t^0 p_X(0) + t^1 p_X(1) + t^2 p_X(2) \dots$

1. $r_X(0) = p_X(0) = P(X=0)$

2. $r_X'(t) = 1p_X(1) + 2t^1 p_X(2) + 3t^2 p_X(3) + \dots$

3. $r_X''(t) = 2p_X(2) + 3 \cdot 2 \cdot t\, p_X(3) + \dots$

$r_X''(0) = 2!\, p_X(2) \Rightarrow p_X(2) = \dfrac{r_X''(0)}{2!} = P(X=2)$

<u>In general</u>

$P(X=k) = p_X(k) = \dfrac{r_X^{(k)}(0)}{k!}$, if you know probability generating function you can generate all probabilities

$X \sim bin(n, \theta)$

$r_X(t) = (1-\theta+\theta t)^n$

$P(X=0) = r_X(0) = (1-\theta)^n$

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

$r_X'(t) = n(1-\theta+\theta t)^{n-1}\theta$

show that $P(X=k) = \binom{n}{k}\theta^k(1-\theta)^{n-k}$

Last hour: Generating function

- p.g.f. $r_x(t) = E[t^x]$
- m.g.f. $m_x(t) = E[e^{tx}]$

connected, if you know one you can get the other

$m_x(t) = r_x(e^t)$

$r_x(t) = m_x(\ln t)$

- Examples:

$X \sim Bin(n, \theta)$

$r_x(t) = (1-\theta+\theta t)^n$
$M_x(t) = (1-\theta+\theta e^t)^n$

$X \sim Geo(\theta)$;

$r_x(t) = \dfrac{\theta}{1-(1-\theta)t}$

$m_x(t) = \dfrac{\theta}{1-(1-\theta)e^t}$

$X \sim P_o(\lambda)$:

$r_x(t) = e^{\lambda(t-1)}$
$M_x(t) = e^{\lambda(e^t-1)}$ $\Big\}\ \forall t \in \mathbb{R}$

$X \sim exp(\lambda) \Rightarrow$

$m_x(t) = \dfrac{\lambda}{\lambda - t}$ $\quad t < \lambda$

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

$M_x(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$ $\quad \forall t \in \mathbb{R}$

$y = a + bx \Rightarrow M_y(t) = e^{at} M_x(bt)$

if x takes values $0, 1, 2, \cdots$, $P(X=k) = \dfrac{r_x^{(k)}(0)}{k!}$

Result

- $M_x^{(k)}(0) = E[X^k]$ ((k): derivative; $E[X^k]$: k^{th} moment of x)

"Pf" $M_x(t) = E[e^{tx}]$

Recall
$M_x(t) = E(e^{tx})$

Pf $M_x(t) = E[e^{tx}]$

$M_x^{(1)}(t) = E[e^{tx} \cdot x] \Rightarrow M_x^{(1)}(0) = E(x)$

$M_x^{(2)}(t) = E[x^2 e^{tx}] \Rightarrow M_x^{(2)}(0) = E(x^2)$

$\vdots$

$M_x^{(k)}(t) = E[x^k e^{tx}] \Rightarrow M_x^{(k)}(0) = E[x^k]$

Ex $x \sim \exp(\lambda) \Rightarrow M_x(t) = \dfrac{\lambda}{\lambda - t}$

$M_x^{(1)} t = \dfrac{\lambda}{(\lambda - t)^2}$ $\quad M_x^{(1)}(O) = \dfrac{\lambda}{\lambda^2} = \dfrac{1}{\lambda} = E(x)$

$M_x^{(2)} t = \dfrac{2\lambda}{(\lambda - t)^3}$ $\quad M^{(2)}_x(0) = \dfrac{2}{\lambda^2} = E(x^2)$

$V(x) = E(x^2) - E(x) = \dfrac{2}{\lambda^2} - \dfrac{1}{\lambda^2} = \dfrac{1}{\lambda^2}$

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.

$M_{x_1+x_2}(t) = E[e^{t(x_1+x_2)}]$

$= E[e^{tx_1} \cdot e^{tx_2}]$

$= E[e^{tx_1}]\, E[e^{tx_2}]$, since $\perp$

$= M_{x_1}(t) + M_{x_2}(t)$

$\blacksquare$

Recall

$M_x(t) = E[e^{tx}]$

In general

$x_1 + x_2 \dots x_n \perp \Rightarrow M_{\sum_{i=1}^{n} x_i}(t) = \prod_{i=1}^{n} M_{x_i}(t)$ | recall $\prod_{i=1}^{n} a_i = a_1 + a_2 \dots a_n$

$r_{\sum x_i}(t) = \prod_{i=1}^{n} r_{x_i}(t)$

Thrm 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. $x \overset{d}{=} y$

So if r.v has $m_x(t) = \frac{\lambda}{\lambda - 1} \Rightarrow x \sim \exp(\lambda)$

NOTE! This does not Apply for mean!

Example Let $x_1, x_2, \dots x_n$ are independent and $x_i \sim N(\mu_i, \sigma^2_i)$

Find the distribution of $y = \sum_{i=1}^{n} x_i$

Find moment generating of y then recognize the distribution

$m_{x_i}(t) = e^{\mu_i t + \frac{1}{2}\sigma_i^2 t^2}$

$m_{\sum x_i}(t) = \prod_{i=1}^{n} m_{x_i}(t)$

$\prod_{1}^{n} (e^{\mu_i t + \frac{1}{2}\sigma_i^2 t^2})$

use result

$m_{\sum_{i=1}^{n} x_i}(t) = \prod_{i=1}^{n} M_{x_i}(t)$

$$= \prod_{i=1}^{n} \left(e^{\mu_i t + \frac{1}{2}\sigma_i^2 t^2} \right)$$

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

$$= M_{\sum x_i}(t) = e^{(\sum_{i=1}^{n} \mu_i)t + \frac{1}{2}(\sum_{i=1}^{n} \sigma^2) t^2}$$

and

$$m_x(t) = e^{\mu + \frac{1}{2}\sigma^2 t^2}$$

$$e^{a_1} e^{a_2} \cdots = \prod e^{a_i} = e^{\sum a_i}$$

By uniqueness of moment generating function

$$\sum_{i=1}^{n} x_i \sim N\left(\sum \mu_i, \sum_{i=1}^{n} \sigma^2\right)$$

But remember for this to be true it has to be normal

Skip 3.5 and get 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

$x \sim Exp(\lambda = 3)$. Find the Markov's upper bound for $p(x \geq \frac{5}{6})$

Answer Markov's upper bound $= \frac{E(x)}{a} = \frac{\left(\frac{1}{3}\right)}{\left(\frac{5}{6}\right)} = \frac{2}{5}$

i.e $p(x \geq \frac{5}{6}) \leq \frac{2}{5}$

(2) Chebyshev's inequality

Let x be any r.v st $V(x) < \infty$ then for $x > 0$

Chebyshev's inequality

Let x be any r.v s.t $\underbrace{V(x) < \infty}_{\text{finite}}$ then for $a > 0$

$$P(|x - \mu_x| \geq a) \leq \frac{V(x)}{a^2}$$

Pf 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

$$P((x - \mu_x)^2 \geq a^2) \leq \frac{E(x - \mu_x)^2}{a^2}$$

$$= P(|x - \mu_x| \geq a) \leq \frac{V(x)}{a^2}$$

Ex $x \sim Exp(\lambda = 3)$. Find the chebyshev's upperbound for $P(|x - \frac{1}{3}| \geq \frac{1}{2})$

a

Ans

$$\text{mean} = \frac{1}{3} = \frac{1}{\lambda}$$

$$\text{Cheby upperbound} = \frac{var(x)}{(\frac{1}{2})^2}$$

$$= \frac{(\frac{1}{\lambda^2})}{(\frac{1}{2})^2} = \frac{\frac{1}{9}}{\frac{1}{4}} = \frac{4}{9}$$

• Cauchy-shewartz inequality

if x & y are any 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

$$\Big(\underbrace{E((x-\mu_x)(y-\mu_y))}\Big)^2 \leq E\left[(x-\mu_x)^2\right] E\left[(y-\mu_y)^2\right]$$

$$\Rightarrow (\mathrm{cov}(x,y))^2 \leq \mathrm{Var}(x)\,\mathrm{var}(y)$$

$$\Rightarrow \frac{(\mathrm{cov}(x,y))^2}{\mathrm{var}(x)\,\mathrm{var}(y)} \leq 1$$

$$\Rightarrow \left(\frac{(\mathrm{cov}(x,y))}{\sqrt{\mathrm{var}(x)\,\mathrm{var}(y)}}\right)^2 \leq 1$$

correlation

$$\mathrm{corr}(x,y)^2 \leq 1$$

$$\Rightarrow -1 \leq \mathrm{corr}(x,y) \leq +1$$

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 of statistics | Let $x_1, x_2 \dots x_n$ be a random sample

① $\bar{x}_n = \dfrac{x_1 + x_2 + \dots x_n}{n} = \sum \dfrac{x_i}{n}$

② $\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$

③ $\max(x_1, \dots x_n)$

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 $\overset{iid}{}$ 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 $\bar{X}_n = \dfrac{\sum_{i=1}^{n} x_i}{n}$

Ans

(a) $x_1 \dots x_n$ independent $x_i \sim N(\mu_i, \sigma_i^2)$

then $\Rightarrow \sum_{i=1}^{n} x_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$

$x_1 \dots x_n \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow T_n \sum_{i=1}^{n} x_i \sim N(n\mu, n\sigma^2)$

b) $T_n \sim N(n\mu, n\sigma^2)$ — μ, σ^2

$M_{T_n}(t) = e^{n\mu t + \frac{1}{2} n \sigma^2 t^2}$

$M_{\bar{x}_n}(t) = e^{n\mu\left(\frac{t}{n}\right) + \frac{1}{2} n \sigma^2 \left(\frac{t}{n}\right)^2}$

$= e^{\mu t + \frac{1}{2} \frac{\sigma^2}{n} t^2}$

Recall

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

$\Rightarrow M_x(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

$y = a + bx$

$M_{\bar{x}_n}(t) = e^{n\mu(\frac{t}{n}) + \frac{1}{2} n\sigma^2 (\frac{t}{n})^2}$

$= e^{\mu t + \frac{1}{2}\frac{\sigma^2}{n} t^2}$

$\Rightarrow \bar{x}_n \sim N(\mu, \frac{\sigma^2}{n})$

$\Rightarrow m_x(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

$y = a + bx$

$a = 0 \quad b = \frac{1}{n}$

$m_y(t) = e^{at} m_x(bt)$

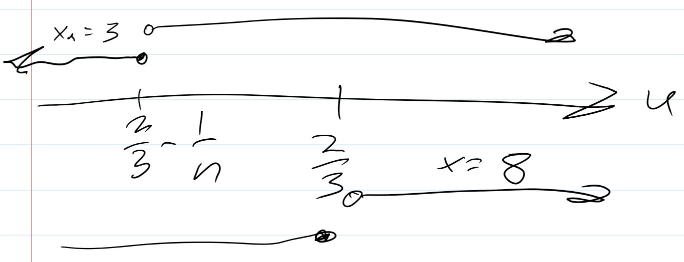
Convergence in Probability

Definition Let $x_1, x_2, \dots$ $\{x_n\}$ be a sequence of random variables then we say x_n converges to x in probability if for any $\epsilon > 0$, $\lim_{n\to\infty} P(|x_n - x| \geq \epsilon) = 0$

$\forall \epsilon > 0, \lim_{n\to\infty} P(|x_n - x| \geq \epsilon) = 0$ / We write $x_n \xrightarrow{P} x$

Ex let $x_n = \begin{cases} 3 & , u \leq \frac{2}{3} - \frac{1}{n} \\ 8 & o/w \end{cases}$

Show that $x_n \xrightarrow{P} x$



$x = 3$

$u \sim Uni(0,1)$

Let $\epsilon > 0$

$P(|x_n - x| \geq \epsilon) = P\left(\frac{2}{3} - \frac{1}{n} < u \leq \frac{2}{3}\right)$

$= \frac{1}{n}$

$0 \leq P(|X_n - X| \geq \varepsilon) = \frac{1}{n}$ <u>squeeze thrm</u>

$\lim P(|X_n - X| \geq \varepsilon) = 0$

i.e $X_n \xrightarrow{P} X$ $\blacksquare$

Last hour / Sampling distributions (4.1)

Convergence

4.2 Convergence in probability

- Definition:

A sequence of random variables $\{X_n\}$ (i.e $X_1, X_2, \ldots$) is said to be convergent to X in probability if

$$\forall \varepsilon > 0 \quad \lim_{n \to \infty} P(|X_n - X| \geq \varepsilon) = 0$$

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, $P(|X - \mu_X| \geq a) \leq \frac{Var(X)}{a^2}$
- Cauchy-Shawartz: 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 $X_1, X_2 \ldots$ be an independent random variables with the same mean

$\Rightarrow E(X_i) = \mu, \ \forall i$

and $V(X_i) \leq v \quad \forall i$ THEN

$$M_n = \frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{P} \mu$$

Converge in probability

pf: $E[M_n] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$

$= \frac{1}{n}\sum_{i=1}^{n} E X_i = \frac{1}{n} \cdot n \cdot \mu = \mu$ ($E X_i = \mu$)

due to linearity

Pf By Chebyshev's inequality

Let $\varepsilon > 0$

$$P[|M_n - \mu| \geq \varepsilon] \leq \frac{Var(M_n)}{\varepsilon^2}$$

$$Var(M_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$\overset{\text{indep!}}{=} \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) \leq \frac{1}{n^2}\sum_{i=1}^{n} v$$

$$= \frac{1}{n^2} \cdot n v$$

$$= \frac{1}{n^2} \cdot nv$$

$$= \frac{v}{n}$$

$$0 \le P[\,|M_n - \mu| \ge \varepsilon\,] \le \frac{Var(M_n)}{\varepsilon^2} \le \frac{v}{n\varepsilon^2} \Rightarrow 0 \text{ as } n \to \infty$$

by squeeze theorem

$$\lim_{n \to \infty} P(|M_n - \mu| \ge \varepsilon) = 0$$

i.e $M_n \xrightarrow{P} \mu$ $\blacksquare$

Today Convergence in distribution

Definition

Let $x_1, x_2, \ldots$ 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

$$\lim_{n \to \infty} P(x_n \le x) = P(X \le x) \quad , \quad \forall x \mid P(X = x) = 0$$

Example

$x_n \sim Ber(\theta = \frac{1}{n})$

F_{x_n}

x	$P_x(x)$
0	$1 - \frac{1}{n}$
1	$\frac{1}{n}$

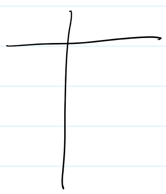
$$F_{x_n}(x) \begin{cases} 0 & , x < 0 \\ 1 - \frac{1}{n} & , 0 \le x < 1 \\ \frac{1}{n} & , x \ge 1 \end{cases}$$

Recall

$$P_x(x) = \begin{cases} \frac{1}{n} & , \\ 1 - \frac{1}{n} & , \end{cases}$$

$$\lim_{n \to \infty} P(x_n \le x) = P(X \le x) \iff \lim_{n \to \infty} F_{x_n}(x) = F_x(x)$$

X:



$$F_X(x) = \begin{cases} 0 & , x < 0 \\ 1 & x \ge 0 \end{cases}$$

Show that $x_n \xrightarrow{D} X \Leftarrow x_n$ converge to x in distribution where

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ 1 & , x \geq 0 \end{cases}$$

Pf $\lim_{n \to \infty} F_{x_n}(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$

$= F_X(x) \; , \forall x$

$\therefore x_n \xrightarrow{D} x$

Central limit Theorem

Let $x_1, x_2 \ldots$ be iid r.v's with finite mean $E x_i = \mu$ and finite variance $Var(x_i) = \sigma^2$, then

$$Z_n = \frac{M_n - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{D.} Z$$ where Z is standard normal distribution

$$M_n = \frac{\sum_{i=1}^{n} x_i}{n}$$

Example

$x_1, x_2, \ldots$ iid with $E(x_i) = \mu = 67$ and $var(x_i) = \sigma^2 = 10^2$

Find the approx probability $P(M_{100} \leq 70)$

$$Z_{100} = \frac{M_{100} - \mu}{\frac{\sigma}{\sqrt{100}}} \xrightarrow{D} Z \quad (CLT)$$

$$= \frac{M_{100} - 67}{\frac{10}{10}} \xrightarrow{D} Z$$

$\frac{10}{10} = 1$

$$P(M_{100} \leq 70) = P\left(\frac{M_{100} - 67}{1} \leq \frac{70 - 67}{1} \right)$$

Closer

closed
not → $\approx P(Z \leq 3) = 1 - P(Z \leq -3)$
exact

$= 0.9987$

use the table

4.2 convergence in prob

~~4.3~~ Conv S ← skipped it, not on exam and 3.5

4.4 convergence in Distribution