MATH 230

Probability and Statistics for Engineers

Spring 2018 (Natalia Zheltukhina)

Lecture Notes

Emin Bahadır Tülüce

Sample Space

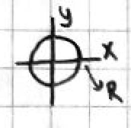
An experiment means an observation which can be reproduced infinitely many times under the same set of conditions.

We will consider idealized experiments.

Def: (*) Any possible outcome of an experiment is called a sample point or simple event.

(**) The set of all sample points is called sample space (Ω or S).

Ex 1) Tossing a coin twice: $\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$

2) Throwing a dart  $\Omega = \{(x,y): x^2+y^2 \leq R^2\}$

Algebra of Events

Def: Let Ω be a sample space. A set of simple events satisfying some set of conditions is called an event (or random event).

Def: An empty set $\emptyset$ is called an impossible event.

The sample space Ω is called a certain event.

Ex Throwing a die, $\Omega = \{1, 2, 3, 4, 5, 6\}$

$A = \{$"At least three scores"$\} = \{3, 4, 5, 6\}$ event

$B = \{$"Eleven scores"$\} = \emptyset$ impossible event

$C = \{$"At most eleven scores"$\} = \Omega$ certain event

$D = \{$"Five scores"$\} = \{5\}$ event (also, simple event)

Def: The event consisting of all sample points that do not belong to A is called "contrary event of A" denoted by $\bar{A}$ or A'.

With any two events A and B we can define new events as;

1) "both A and B occur" $\longrightarrow A \cap B$
2) "either A or B (or both) occur" $\longrightarrow A \cup B$

Def: If $A \cap B = \emptyset$ then A and B are called "mutually exclusive".

De Morgan's Formula

$$\overline{A \cap B} = \bar{A} \cup \bar{B} \qquad \overline{A \cup B} = \bar{A} \cap \bar{B} \xrightarrow{\text{generalize}} \overline{\bigcap_i A_i} = \bigcup_i \bar{A}_i \qquad \overline{\bigcup_i A_i} = \bigcap_i \bar{A}_i$$

Def: (Classical Definition of Probability): Let Ω be a finite sample space consisting n sample points, $\Omega = \{w_1, w_2, \dots, w_n\}$. Let $A \subset \Omega$ be a random event. The probability of A is,

$$P(A) = \frac{|A|}{|\Omega|}$$ where $| \, |$ indicates the number of elements.

Ex Tossing a coin twice. $\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$

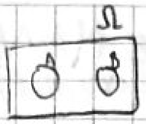
$A = \{\text{"just one tail"}\} = \{(H,T), (T,H)\}$ $\qquad P(A) = \frac{2}{4} = \frac{1}{2}$

Properties of Classical Probability

(1) $0 \leq P(A) \leq 1$ for any random event A. $\qquad$ (2) $P(\Omega) = 1$

(3) If A and B are mutually exclusive; $P(A \cup B) = P(A) + P(B)$

(4) $P(\bar{A}) = 1 - P(A)$ $\qquad$ (5) If $A \subset B$ then, $P(A) \leq P(B)$

Proof of (3):  $P(A \cup B) = \frac{|A \cup B|}{|\Omega|} = \frac{|A| + |B|}{|\Omega|} = \frac{|A|}{|\Omega|} + \frac{|B|}{|\Omega|} = P(A) + P(B)$

Proof of (4): $\underbrace{1 = P(\Omega)}_{(1)} \underbrace{=}_{\Omega = A \cup \bar{A}} P(A \cup \bar{A}) \underbrace{=}_{(3)} P(A) + P(\bar{A})$

Axiomatic Approach to Probability

Let Ω be a sample space of some experiment.

Def: The probability of a random event is a real-valued function from the set of all subsets of Ω to $[0, 1]$, satisfying followings:

(1) $P(A) \geq 0$, $\forall A \subset A$ (Axiom 1)

(2) $P(\Omega) = 1$ (Axiom 2)

(3) $P(\bigcup_{j=1}^{\infty} A) = \sum_{j=1}^{\infty} P(A_j)$ where $A_1, A_2 \dots$ are mutually disjoint. (Axiom 3)

that is $A_i \cap A_j = \emptyset$ for any $i \neq j$.

Properties of Probability

(1^*) $P(\emptyset) = 0$ Proof: Take all $A_j = \emptyset$, $\bigcup_{j=1}^{\infty} A_j = \emptyset \Rightarrow P(\emptyset) \overset{(3)}{=} P(\emptyset) + P(\emptyset) + \dots + P(\emptyset) \Rightarrow P(\emptyset) = 0$

(2^*) For any finite sequence of n mutually exclusive events $A_1, A_2, \dots, A_n$.

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_3)$$

Proof: Take $A_{n+1}, A_{n+2}, \dots$ as $\emptyset$ in Axiom 3.

(3^*) $P(A) = 1 - P(\bar{A})$ Proof: $1 \overset{(1)}{=} P(\Omega) \overset{\Omega = A \cup \bar{A}}{=} P(A \cup \bar{A}) \overset{?}{=} P(A) + P(\bar{A})$

Corollary: $0 \leq P(A) \leq 1$ → From (3^*) $\Leftrightarrow P(A) = 1 - P(\bar{A})$

(4^*) If $A \subset B$ then $P(A) \leq P(B)$

(5^*) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(6*) $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + \dots + (-1)^n P(A_1 \cap A_2 \cap \dots \cap A_n)$

Uniform Space

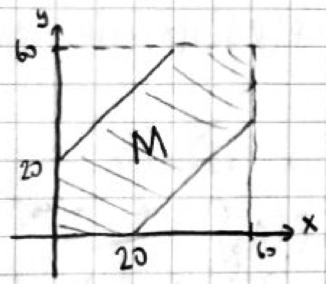
Let the sample space Ω be a bounded region in $\mathbb{R}^n$. $\begin{pmatrix} \Omega \subset \mathbb{R} & \text{(line)} \\ \Omega \subset \mathbb{R}^2 & \text{(area)} \end{pmatrix}$

The probability of event A is, $P(A) = \dfrac{mes(A)}{mes(\Omega)}$ where $mes(A)$ means either length / area / volume according to its dimension.

Ex) Two person A and B are going to go to the meeting place between a random time in between 12:00 - 13:00. After arriving, they wait 20 minutes and leave. What is the probability of they will see each other?

$\Omega = \{(x, y) : 0 \leqslant x \leqslant 60 \ , \ 0 \leqslant y \leqslant 60\}$

$M = \{\text{"A and B will meet."}\} = \{(x, y) : |x - y| \leqslant 20\}$



$M = \{(x, y) : -20 \leqslant x - y \leqslant 20\}$

$P(M) = \dfrac{Area(M)}{Area(\Omega)} = \dfrac{60^2 - \frac{1}{2} 40^2 \cdot 2}{60^2} = 1 - \dfrac{40^2}{60^2} = \dfrac{5}{9}$

Finite Sample Space

Let $\Omega = \{w_1, w_2, \dots, w_n\}$. If to any sample point w_i, there corresponds a real number $p_i \geqslant 0$ and $p_1 + p_2 + \dots + p_n = 1$ then the function

$P(A) = \sum_{w_i \in A} p_i$ defines the probability function.

Ex) $\Omega = \{w_1, w_2, w_3\}$ with $p_1 = \frac{1}{6}$, $p_2 = \frac{1}{6}$, $p_3 = \frac{2}{3}$ $\Rightarrow$

$P(\{w_1, w_2\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

$P(\{w_2, w_3\}) = \frac{1}{6} + \frac{2}{3} = \frac{5}{6}$

Ex) An urn contains 1000 lottery tickets numbered 1 to 1000. One is selected at random by paying \$2. You will obtain \$3 if the number is divisible by 2, 3 or 5. You lose your \$2 if not. Will you play?

Denote D_k, the probability of the number is divisible by k.

int part

$P(D_2) = \frac{500}{1000} = \frac{1}{2}$

$P(D_3) = \frac{[1000/3]}{1000} = \frac{333}{1000}$

$P(D_5) = \frac{200}{1000} = \frac{1}{5}$

$$P(D_2 \cup D_3 \cup D_5) = P(D_2) + P(D_3) + P(D_5) \quad (+)$$
$$- P(D_6) - P(D_{10}) - P(D_{15}) \quad (-)$$
$$+ P(D_{30}) \quad (+) \quad = P(W)$$

$P(D_6) = \frac{166}{1000}$, $P(D_{10}) = \frac{100}{1000}$, $P(D_{15}) = \frac{66}{1000}$, $P(D_{30}) = \frac{33}{1000}$

$$P(W) = \frac{500 + 333 + 200 - 166 - 100 - 66 + 33}{1000} = \frac{367}{500} > \frac{1}{2} \Rightarrow$$

Combinatorial Analysis

Basic Principle of Counting: If r many experiments are performed which has $n_1, n_2, \ldots, n_r$ outcomes, then, there is a total of $n_1 n_2 \cdots n_r$ many outcomes possible.

Def: Any arrangement of r objects, taken from n objects is called r-permutation.

Theorem: i. The number of permutations of n distinct objects is $n!$

ii. The number of r-permutations of n distinct objects is $\frac{n!}{(n-r)!}$

Proof: Any r-permutation is a result of r experiments:

Experiment 1: Choose an object out of n to put #1 (n outcomes)

Experiment 2: Choose an object out of $(n-1)$ to put #2 $(n-1$ outcomes$)$

Experiment 3: " " " $(n-2)$ to put #3 $(n-2$ outcomes$)$

Experiment r: " " " $(n-r+1)$ to put #r $(n-r+1$ outcomes$)$

$$\#\text{outcomes} = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}$$

Ex) How many different arrangements are there of word PEPPER ?

2E, 3P are not distinct (distinguishable). $\Rightarrow = \frac{6!}{3!\,2!} = 60$

Theorem: If there are n objects with n_1 of a 1^{st} type, n_2 of a 2^{nd} type, ..., n_r of r^{th} type, where $n_1 + n_2 + n_3 + \cdots + n_r = n$ then, there are $\frac{n!}{n_1!\cdot n_2!\cdots n_r!}$ arrangements of given objects.

$\frac{P(n,r)}{r!}$

Theorem: $\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$ Proof: Each r-combination produces $r!$ r-permutations

Theorem: (Binomial) $$(x+y)^n = \sum_{k=0}^{n} C(n,k)\, x^k y^{n-k}$$

Divide n distinct elements into r (numbered) groups with sizes $n_1, n_2, \dots, n_r$ where $n_1+n_2+\dots+n_r = n$. How many different divisions are possible?

Answer: $$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n_r}{n_r} = \cdots = \frac{n!}{n_1!\cdot n_2!\cdot n_3!\cdots n_r!}$$

Notation: $$\frac{n!}{n_1!\cdot n_2!\cdots n_r!} = \binom{n}{n_1, n_2, \dots, n_r}$$

Ex) 10 childrens will form team A and team B with 5 childrens in it. How many ways are possible?

$$\left(\binom{10}{5}\right)\binom{10}{5,5} = \frac{10!}{5!\cdot 5!}$$

Ex| How many different ways 10 children can make a 5v5 match?

$\frac{1}{2!}\binom{10}{5,5} = \frac{10!}{5!\cdot 5!\cdot 2!}$ ← Teams are not numbered.

Multinomial Theorem: $(x_1 + x_2 + \dots + x_r)^n = \sum_{\substack{n_1+n_2+\dots+n_r=n \\ 0\le n_i\le n}} \binom{n}{n_1, n_2, \dots n_r} \cdot x_1^{n_1} x_2^{n_2} \dots x_r^{n_r}$

Ex| $(x_1+x_2+x_3)^2 = \binom{2}{2,0,0}x_1^2 + \binom{2}{0,2,0}x_2^2 + \binom{2}{0,0,2}x_3^2$

$2 = 2+0+0$
$= 0+2+0$
$= 0+0+2$
$= 1+1+0$
$= 1+0+1$
$= 0+1+1$

$+ \binom{2}{1,1,0}x_1x_2 + \binom{2}{1,0,1}x_1x_3 + \binom{2}{0,1,1}x_2x_3$

$= x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$

→ Numbers $\binom{n}{n_1, n_2, \dots, n_r}$ are called multinomial coefficients.

Conditional Probability: $P(X \mid Y)$: prob of x, given Y.

Ex| Probability of royal flush for 5 taken cards (same suit of J, Q, K, A, 10): $\frac{4}{\binom{52}{5}}$ ← 4 different royal flushes

Suppose that we know our first card is K of spades: $\frac{1}{\binom{51}{4}}$ → royal flush of spades

Def: Let Ω be a sample space, A and B be events where $P(B) \neq 0$.
Given that B occurs, the conditional probability that A occurs is denoted by

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

$\Rightarrow P(A \cap B) = P(B) \cdot P(A \mid B) \Rightarrow \dots \Rightarrow$

$$P(A) = P(A \mid B)\,P(B) + P(A \mid \bar{B})\,P(\bar{B})$$

Ex) In answering a question on a multiple-choice test, a student either knows the answer or guesses. Let p be the probability that student knows the answer. Assume that the student guesses correct with prob. $\frac{1}{m}$ where m is the number of alternatives. What is the probability that a student know the answer given that he answered correctly.

$C = \{\text{Student answers correctly}\}$

$K = \{\text{Student knows the answer}\}$

Given

$P(K) = p$

$P(C \mid \bar{K}) = \frac{1}{m}$

Aim

$P(K \mid C) = ?$

$$P(K \mid C) = \frac{P(K \cap C)}{P(C)} = \frac{P(K) \cdot P(C \mid K)}{P(K \cap C) + P(\bar{K} \cap C)} = \frac{P(K) \cdot P(C \mid K)}{P(K)P(C \mid K) + P(\bar{K})P(C \mid \bar{K})}$$

$$= \frac{p \cdot 1}{p \cdot 1 + (1-p) \cdot \frac{1}{m}} = \frac{mp}{mp + (1-p)}$$

Theorem: Let $A \subset \bigcup_{i=1}^{n} B_i$ and $B_i \cap B_j \neq \emptyset$, $i \neq j$

Then, $P(A) = P(A \mid B_1) P(B_1) + P(A \mid B_2) P(B_2) + \dots + P(A \mid B_n) P(B_n)$

$$P(A) = \sum_{k=1}^{n} P(A \mid B_k) \cdot P(B_k)$$

Quiz

Corollary (Baye's Theorem): If $B_1, B_2, \dots, B_n$ form a partition of Ω, then, any random event A;

$$P(B_i \mid A) = \frac{P(A \mid B_i) P(B_i)}{\underbrace{\sum_{k=1}^{n} P(A \mid B_k) P(B_k)}_{P(A)}}$$

Def: Events A and B are said to be independent if $P(A\cap B)=P(A)\cdot P(B)$

Remark: A and B independent $\iff P(A\mid B)=\dfrac{P(A\cap B)}{P(B)}=\dfrac{P(A)\,P(B)}{P(B)}=P(A)$

Ex Throwing a die $\quad A=\{\text{even}\}=\{2,4,6\}$

$B=\{1,2,3,4\}$ $\qquad$ Are A and B independent?

$$\left.\begin{aligned}&P(A\cap B)=P(\{2,4\})=\tfrac{1}{3}\\&P(A)=\tfrac{1}{2},\quad P(B)=\tfrac{2}{3}\Rightarrow P(A)\cdot P(B)=\tfrac{1}{3}\end{aligned}\right\}$$ They are independent.

If $B=\{1,2,3,4,5\}$, they would not be.

Theorem: If A and B are independent;

i. $\overline{A}$ and B are independent.

ii. A and $\overline{B}$ are independent.

iii. $\overline{A}$ and $\overline{B}$ are independent.

Remark: If $P(A)=0$, then A and B are independent for all B.

Def: Events $A_1, A_2, \dots, A_n$ are independent if for any collection $A_{k1}, A_{k2}, \dots, A_{kr}$, $P(A_{k_1}\cap A_{k2}\cap\dots A_{kr})=P(A_{k1})\cdot P(A_{k_2})\cdots P(A_{kr})$.

For example, three events A_1, A_2, A_3 are independent if,

$P(A_1\cap A_2)=P(A_1)P(A_2)$

$P(A_2\cap A_3)=P(A_2)P(A_3)$, $\quad P(A_1\cap A_2\cap A_3)=P(A_1)P(A_2)P(A_3)$

$P(A_1\cap A_3)=P(A_1)P(A_3)$

are all satisfied.

Theorem: If $A_1, A_2, \dots, A_n$ are independent, then $\overline{A}_1, \overline{A}_2, \dots \overline{A}_n$ are independent.

Theorem: If $A_1, A_2, \dots, A_n$ are independent, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - P(\bar{A}_1)P(\bar{A}_2)\cdots P(\bar{A}_3)$$

Ex) We have 3 cards with sides of color: RR, BB, RB. One card is selected and has a color of RED on its side. What is prob. of the other side being BLACK?

$$P\left(RB \mid \underset{R}{\text{upper}}\right) = \frac{P\left(\underset{R}{\text{upper}} \mid RB\right)\cdot P(RB)}{P\left(\underset{R}{\text{upper}} \mid RB\right)\cdot P(RB) + P\left(\underset{R}{\text{upper}} \mid RR\right)\cdot P(RR) + P\left(\underset{R}{\text{upper}} \mid BB\right)\cdot P(BB)} = \frac{1}{3}$$

The number of integer non-negative solutions to equation

$$X_1 + X_2 + \cdots + X_n = r \quad , \quad \boxed{\text{Sols} = \binom{r+n-1}{n-1}}$$

→ $x_1 + x_2 + x_3 = 5$

$2+1+2 \iff 0010100$ $\quad$ $1+1+3 \iff 0101000$
$0+0+5 \iff 1100000$ $\quad$ $2+0+3 \iff 0011000$

$\}$ Number of arrangements $= \binom{r+n-1}{n-1} = \binom{5+3-1}{3-1} =$

Ex) How many ways are there to distribute 2 white and 20 red balls into 10 distinct boxes, so that exactly one box is empty.

$$\begin{pmatrix}\text{whites are in} \\ \text{same box}\end{pmatrix} + \underbrace{\begin{pmatrix}\text{whites are in} \\ \text{diff. boxes}\end{pmatrix}}_{X} = \begin{pmatrix}\text{choose} \\ \text{empty} \\ \text{box}\end{pmatrix}\cdot\begin{pmatrix}\text{Choose} \\ \text{box for} \\ \text{whites}\end{pmatrix}\cdot\begin{pmatrix}\text{place} \\ \text{remaining}\end{pmatrix} + X$$

put 1 red on each empty box

$$10 \cdot 9 \cdot \binom{12+8}{8}$$

$$X = \begin{pmatrix}\text{Choose} \\ \text{empty} \\ \text{box}\end{pmatrix}\begin{pmatrix}\text{choose} \\ \text{2 boxes} \\ \text{for white}\end{pmatrix}\cdot\begin{pmatrix}\text{place} \\ \text{remaining}\end{pmatrix}$$

put 1 red on each empty

$$10 \cdot \binom{9}{2} \cdot \binom{13+8}{8}$$

ADD

\# Reminders: $P(A\cap B) = P(A \mid B)\cdot P(B)$ ⇐ $P(A|B) = \dfrac{P(A\cap B)}{P(B)}$

$$P(B_i \mid A) = \frac{P(B_i \cap A)}{\sum P(B_k \cap A)} = \frac{P(A\cap B_i)\,P(B_i)}{\sum P(A\cap B_k)\,P(B_k)}$$ where $B_1, B_2 \ldots$ are partitions of Ω

Random Variables

Real valued functions $X: \Omega \to \mathbb{R}$ are called random variables.

Ex) Tossing a fair coin 3 times. X = # heads appeared

X takes $0, 1, 2, 3$.

$P(X=0) = 1/8$ $\quad P(X=2) = 3/8$

$P(X=1) = 3/8$ $\quad P(X=3) = 1/8$

Probability dist. of rand. var. X:

X	0	1	2	3
P	1/8	3/8	3/8	1/8

Ex) Tossing a coin until a head. X = # tossing

X	1	2	3	...
P	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	...

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} \cdots = \frac{a}{1-r} = \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1$$

Def: X is discrete if it takes finite number of value or countably infinite number of values.

Ex) We took 3 different balls of 6 balls with 1-6 numbers on it.

X: The largest number selected. Find the probability distribution of X.

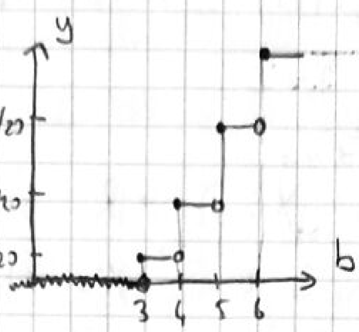
$P(X=3) = \dfrac{1}{\binom{6}{3}} = \dfrac{1}{20}$ $\qquad P(X=5) = \dfrac{\binom{4}{2}}{\binom{6}{3}} = \dfrac{6}{20}$

$P(X=4) = \dfrac{\binom{3}{2}}{\binom{6}{3}} = \dfrac{3}{20}$ $\qquad P(X=6) = \dfrac{\binom{5}{2}}{\binom{6}{3}} = \dfrac{10}{20}$

Def: The (cumulative) distribution function $F(b)$ for a random variable X is defined as $F(b) = P(X \leq b)$, $F: \mathbb{R} \to \mathbb{R}$

Ex In prev. example

$$F(b) = \begin{cases} 0 , & b < 3 \\ \frac{1}{20} , & 3 \leq b < 4 \\ \frac{4}{20} , & 4 \leq b < 5 \\ 1 , & 6 \leq b \end{cases}$$



$F(5) = 10/20$

$F(5-0) = 4/20$

Properties: (1) $F(b)$ is non decreasing.

(2) $\lim_{b \to -\infty} F(b) = 0$ (3) $\lim_{b \to \infty} F(b) = 1$

(4) F is right-continuous: $\lim_{b \to a^+} F(b) = F(a)$.

Notation: $F(a-0) = \lim_{b \to a^-} F(b)$

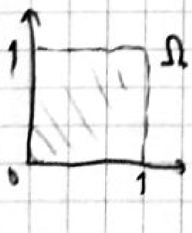
Ex (i) $P(a < X \leq b) = F(b) - F(a)$

(ii) $P(a \leq X \leq b) = F(b) - F(a-0)$

(iii) $P(a < X < b) = F(b-0) - F(a)$

(iv) $P(X < a) = 1 - F(a)$

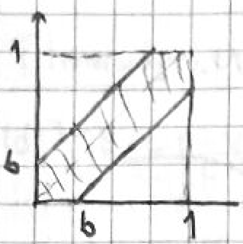
Ex Take a point on square $[0,1] \times [0,1]$ at random. Find the distribution function of the random variable $X = |x-y|$.



$F(b) = P(X \leq b)$

If $b < 0 \Rightarrow F(b) = 0$

If $b \geq 1 \Rightarrow F(b) = 1$



If $0 \le b < 1$, then $F(b) = P(X \le b)$

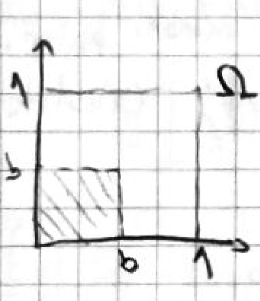
$$P(|x-y| \le b) = \frac{\text{Area}(|x-y| \le b)}{\text{Area}(\Omega)} = \frac{1 - 2 \cdot \frac{1}{2}(1-b)^2}{1} = 2b - b^2$$

$$F(b) = \begin{cases} 0 & b < 0 \\ 2b - b^2 & 0 \le b < 1 \\ 1 & 1 \le b \end{cases}$$

Ex) Take a point on square $[0,1] \times [0,1]$. Find the distribution of random variable $Y = \max(x, y)$.

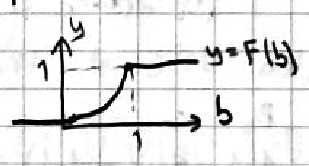
If $b < 0$, then $F(b) = 0$
If $b \geq 1$, then $F(b) = 1$

If $0 \le b < 1$ then $F(b) = P(X \le b)$



$$F(b) = P(\max(x,y) \le b) = \frac{b^2}{1} = b^2$$

$$F(b) = \begin{cases} 0 & b < 1 \\ b^2 & 0 \le b < 1 \\ 1 & 1 \le b \end{cases}$$



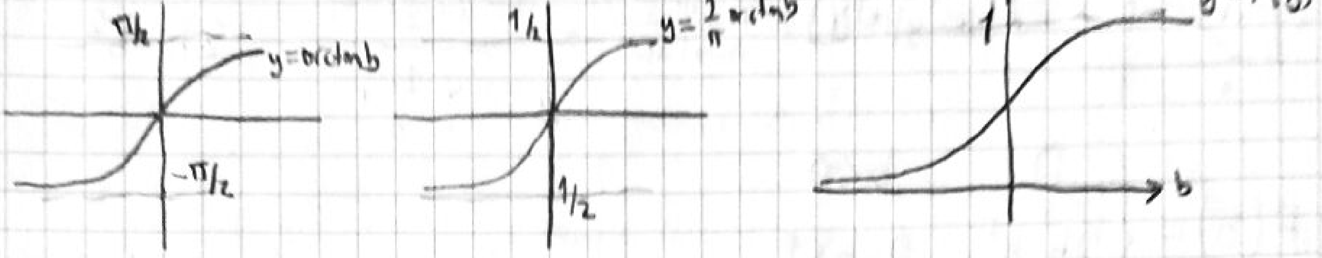
Ex) Let $F(b) = A + B\arctan b$ be a distribution of rand. var. X.

(a) Find constants A and B.
(b) Sketch the graph of $F(b)$.
(c) Find the probability $P(-1 \le X \le 1)$

(a) $\lim\limits_{b \to \infty} A + B\arctan b = 1 \Rightarrow A + B(\pi/2) = 1$ $\quad A = 1/2$

$\lim\limits_{b \to \infty} A + B\arctan b = 0 \Rightarrow A + B(-\pi/2) = 0$ $\quad B = 1/\pi$

(b) $y = F(b) = \frac{1}{2} + \frac{1}{\pi} \arctan b$



(c) $P(-1 \leqslant X \leqslant 1) = P(X \leq 1) - P(X < -1)$

$$= F(1) - F(-1-0) = \left(\frac{1}{2} + \frac{1}{\pi}\cdot\frac{\pi}{4}\right) - \left(\frac{1}{2} + \frac{1}{\pi}\left(\frac{-\pi}{4}\right)\right) = \frac{1}{2}$$

If X takes $a_1, a_2, \dots, a_N$ where $N \leqslant \infty$,

Def: The expected value (expectation) of discrete random variable X is

$$E[X] = a_1 P(X = a_1) + a_2 P(X = a_2) + \cdots = \sum_{k=1}^{N} a_k \cdot P(X = a_k)$$

Ex | A box contains 5 red and 5 blue marbles. Two marbles picked. If they are the same, you win \$1.10, otherwise lose \$1. Calculate the expected value of the amount you gain (or lose).

$$E[X] = 1.1 \cdot P(X = 1.1) + (-1) \cdot P(X = -1) = ?$$

$|\Omega| = \binom{10}{2} = 45$

X	−1	1.1
p	5/9	4/9

$P(X=-1) = \frac{\binom{5}{1}\binom{5}{1}}{\binom{10}{2}}$, $P(X = 1.1) = \frac{\binom{5}{2} + \binom{5}{2}}{\binom{10}{2}}$

$$E[X] = 1.1 \cdot \frac{4}{9} + (-1)\cdot\frac{5}{9} = \frac{4.4 - 5}{9} = -\frac{6}{90} = -\frac{1}{15}$$ dollars

Theorem: Let X be a discrete rand. var. Let c, d be constants.

(a) $E[cX+d] = cE[X]+d$

(b) If X takes the value b maximum, $E[X] \leq b$.

Ex) Let X be a rand. var. that takes $(-1, 0, 1)$ with probs $(0.2, 0.5, 0.3)$

Compute the value $E[X^2]$

X^2	0	1
P	1/2	1/2

$P(X^2=0) = P(X=0) = 0.5$

$P(X^2=1) = [P(X=-1) + P(X=1)] = 0.2 + 0.3 = 0.5$

$E[X^2] = 0 \cdot 0.5 + 1 \cdot 0.5 = \underline{\underline{0.5}}$

Theorem: Let X be disc. rand. var. with values $a_1, a_2, \ldots$ then, for any real function g, we have;

$$E[g(X)] = \sum_{k=1}^{N} g(a_k) \cdot P(X = a_k) \quad \Big\} \ (*)$$

Ex) For the above question; $E[X^2] = (-1)^2 \cdot 0.2 + 0^2 \cdot 0.5 + 1^2 \cdot 0.3 = \underline{\underline{0.5}}$

Def: The variance of X is defined as;

$$V(X) = E\left[(X - E[X])^2\right]$$

standard dev. of X, $\sigma = \sqrt{V(X)}$

Ex) For the above question

$$V(X) = E\left[(X-0.1)^2\right] = [(-1)-0.1]^2 \cdot 0.2 + [0-0.1]^2 \cdot 0.5 + [1-0.1]^2 \cdot 0.3$$

$$= 1.21 \cdot 0.2 + 0.01 \cdot 0.5 + 0.81 \cdot 0.3 = \ldots$$

Ex For two coin flips, you win \$1 if its two tails
you win \$2 if its two heads
you lose \$1 if its heads and tails.

X= (amount of money we win) $\Omega = \{HH, TT, \underbrace{HT, TH}_{(-1)}\}$ with HH (+2), TT (+1)

X	-1	1	2
P	1/2	1/4	1/4

$$E[x] = (-1)\frac{1}{2} + 1\frac{1}{4} + 2\frac{1}{4} = 1/4$$

$$V(x) = E\left[\left(X-\tfrac{1}{4}\right)^2\right] = \left(-1-\tfrac{1}{4}\right)^2\cdot\frac{1}{2} + \left(1-\tfrac{1}{4}\right)^2\cdot\frac{1}{4} + \left(2-\tfrac{1}{4}\right)^2\cdot\frac{1}{4} =$$

$$V(x) = E\left[(X-E[X])^2\right] = \sum_{k=1}^{N}\left(a_k - E[X]^2\right)P(X=a_k)$$

$$= \sum_{k=1}^{N}\left(a_k^2 - 2E[x]a_k + E[x]^2\right)P(X=a_k)$$

$$= \left(\sum_{k=1}^{N} a_k^2 P(X=a_k)\right) - \left(2E[x]\cdot\sum_{k=1}^{N} a_k P(X=a_k)\right) + E[x]^2\cdot\sum_{k=1}^{N} P(X=a_k)$$

$$= E[X^2] - 2E[X]^2 + E[X]^2$$

$$\boxed{V(X) = E[X^2] - E[X]^2} \rightarrow \text{computational formula for } V(X).$$

Ex Calculate the variance of X = (a die rolled once)

$$E(X) = \sum_{k=1}^{6} a_k\cdot P(X=a_k) = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6}\cdots 6\cdot\frac{1}{6} = \frac{21}{6} = 3.5$$

$$V(x) = E\left[(X-3.5)^2\right] = \sum_{k=1}^{6}(a_k-3.5)^2\cdot P(X=a_k) = \frac{1}{6}\left(2.5^2 + 1.5^2 + 0.5^2 + 0.5^2 + 1.5^2 + 2.5^2\right) = \frac{35}{12}$$

$$V(x) = E[x^2] - E[X]^2 = \frac{1}{6}\left(1+2^2+3^2\ldots 6^2\right) - 3.5^2 = \frac{35}{12}$$

Theorem: $V(aX+b) = a^2 \cdot V(X)$ Proof: Write $V(aX+b) = \cdots$

Ex $E[X]=1$, $V(X)=5$, $E[(2+X)^2]=?$

$$E[(2+X)^2] = E[X^2+4X+4] = E[X^2]+4E[X]+4$$

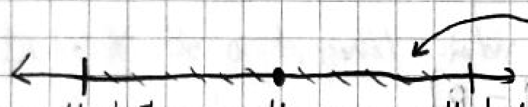
$$= \underbrace{E[X^2]-E[X]^2}+E[X]^2+4E[X]+4$$

$$= V[X] + E[X]^2 + 4E[X] + 4 = 5 + 1^2 + 4\cdot 1 + 4 = 14$$

Chebyshev Theorem: Let X be a rand var. with expected value $E[X]=\mu$ and standard deviation σ. Then,

$$P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2} \quad \text{for any } k>1.$$

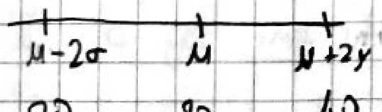
Explanation:



we have $\geq \left(1-\frac{1}{k^2}\right)N$ of all values taken by X.

Ex Let X take $N=200$ values with $\mu = E[X] = 30$, $\sigma = \sqrt{V(X)} = 5$

Take $k=2$ and apply Chebyshew theorem.



$$\Rightarrow P(20 < X < 40) \geq 1 - \frac{1}{2^2} = 0.75$$

Ex] In a factory, daily production of motors is averaged 120, with st. deviation 10.
(a) What fraction of days will have production between 100 – 140
(b) Find the shortest interval certain to contain at least 90% of the daily production

$X = $ (amount of motors produced daily)

(a) $E[X] = 120 \quad \sigma = 10$, take $k = \frac{140-120}{10} = 2$

By Chebyshev theorem, $\geq 1 - \frac{1}{2^2} = \frac{3}{4}$ of all values are in $(100, 140)$.

(b) $1 - \frac{1}{k^2} = 0.9 \Rightarrow k = \sqrt{10} \Rightarrow k\sigma = 10\sqrt{10} \Rightarrow (120 - 10\sqrt{10}, 120 + 10\sqrt{10})$

Bernoulli Random Variable: X is called Bernoulli rand. var. if it takes only two values 0 and 1. 1 for success, 0 for failure

X	0	1
P	1-p	p

$E[x] = p$
$V(x) = p(1-p)$

Binomial Random Variable: The binomial experiment;

1- The same trial is repeated n times.
2- Each trial has two outcomes: succ / fail
3- Probability of success in each trial stays the same, say p.
4- Trials are independent.

Let $X = $ (number of successes in this binomial experiment.)

X is called a binomial random variable with parameters n and p.

X	0	1	2	...	K	...	n
P	$(1-p)^n$	$\binom{n}{1}p^1(1-p)^{n-1}$	$\binom{n}{2}p^2(1-p)^{n-2}$	...	$\binom{n}{K}p^K\cdot(1-p)^{n-K}$	...	p^n

$$\sum_{K=0}^{n} P(X=K) = \sum_{K=0}^{n} \binom{n}{K} p^K (1-p)^{n-K} = (p+(1-p))^n = 1$$

Ex| A student takes a test with 5 multiple choice questions with 5 choices each. Find the probability that he will answer exactly 3 questions correctly.

$p = \frac{1}{4}$ $X = \#$ correct answers → binomial rand. var, $n=5$.

$1-p = \frac{3}{4}$

$P(X=3) = ?$

$$P(X=3) = \binom{5}{3}\left(\frac{1}{4}\right)^3 \cdot \left(\frac{3}{4}\right)^2 = \frac{90}{4^5}$$

Theorem: Let X be a binomial number with parameters n and p. Then,

$$E[X] = np \qquad V(X) = np(1-p)$$

Geometric Random Variable: Independent identical Bernoulli trials are performed until a success occurs. Let X be the number of trials.

X	1	2	3	...	K
P	p	$p(1-p)^1$	$p(1-p)^2$	...	$p(1-p)^{K-1}$

$$\sum_{K=1}^{\infty} P(X=K) = \sum_{K=1}^{\infty} (1-p)^{K-1}\cdot p = p \cdot \sum_{j=0}^{\infty} (1-p)^j = p \cdot \frac{1}{1-(1-p)} = 1$$

The random variable X is called geometric with parameter p.

Ex] A recuiting firm finds that 30% of applicants have training. Find the probability that first occurence of an applicant of trained is found on fifth interview.

$X =$ # trials to find a trained applicant

$X =$ Geometric with $0.3 \Rightarrow P(X=5) = (0.7)^4\, 0.3$

Theorem: Let X be a geometric random variable with parameter p.

$$E[X] = \frac{1}{p} \quad , \quad V(X) = \frac{1-p}{p^2}$$

The Moment-Generating Function: Let Y be a random variable. The moment generating function of Y is defined as

$$\boxed{M(t) = E[e^{tY}]}$$

then, $M(t) = E[e^{tY}] = \sum_{k=1}^{N} e^{ty_k} \cdot P(Y=y_k)$. We have;

$$M'(t) = \sum_{k=1}^{N} y_k e^{ty_k} \cdot P(Y=y_k) = E[Ye^{tY}] \quad , \quad M'(0) = E[Y] \Big\} \text{ 1st moment of } Y.$$

$$M''(t) = \quad = E[Y^2 \cdot e^{tY}] \quad , \quad M''(0) = E[Y^2] \Big\} \text{ 2nd moment of } Y.$$

$$\vdots$$

$$M^{(n)}(t) = \quad = E[Y^n \cdot e^{tY}] \quad , \quad M^{(n)}(0) = E[Y^n]$$

Ex] Use moment-generating to find the expected value and variance for the Binomial random variable X with parameters n and p.

$$M(t) = E[e^{tX}] = \sum_{k=0}^{N} e^{tk} \cdot P(X=k) = \sum_{k=0}^{N} e^{tk} \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} \underbrace{(pe^t)^k}_{a^k} \underbrace{(1-p)^{n-k}}_{b^{n-k}} = (pe^t + 1 - p)^n \Rightarrow \begin{array}{l} M'(t) = n(pe^t + 1 - p)^{n-1} \cdot pe^t \\ M'(0) = np = E[X] \end{array}$$

$$V[X] = E[X^2] - E[X]^2 = M''(0) - (M'(0))^2$$

$$M''(t) = n(n-1)(pe^t + 1-p)^{n-2}\, pe^t \cdot pe^t + n(pe^t + 1-p)^{n-1} \cdot pe^t$$

$$M''(0) = n(n-1)\,p^2 + np = n^2p^2 - np^2 + np$$

$$V[X] = n^2p^2 - np^2 + np - (np)^2 = \underline{np(1-p)}$$

The Negative Binomial Random Variable: Independent Bernoulli trials are performed until a total of r successes is accumulated. Let X count the number of required trials.

X	r	$r+1$		n	
p	p^r	$\binom{r}{r-1} p^r (1-p)$	...	$\binom{n-1}{r-1} p^r (1-p)^{n-r}$	...

X is called negative Binomial Random Variable with parameters r, p.

Note that $\sum_{n=r}^{\infty} \binom{n-1}{r-1} p^r (1-p)^{n-r} = 1$

$$E[X] = \sum_{n=r}^{\infty} n \binom{n-1}{r-1} p^r (1-p)^r = \cdots = \frac{r}{p} \cdot \sum_{n+1=r+1}^{\infty} \binom{(n+1)-1}{(r+1)-1} p^{r+1} (1-p)^{(n+1)-(r+1)} = \frac{r}{p} \cdot 1$$

Theorem: (1) For a geometric random variable X with parameter p;

$$P(X=k) = (1-p)^{k-1} p \quad , \quad E[X] = \frac{1}{p} \quad , \quad V(X) = \frac{1-p}{p^2}$$

(2) For a negative binomial random variable X with parameters r, p;

$$P(X=k) = \binom{k-1}{r-1} (1-p)^{k-r} p^r$$

$$E[X] = \frac{r}{p} \quad , \quad V(X) = \frac{r(1-p)}{p^2}$$

Hypergeometric Random Variable: We have an urn with m white and $N-m$ black balls. We choose n balls without replacement. Let X count the number of white balls in our selection.
X is called hypergeometric random variable with params N, m, n.

$$P(X=K) = \frac{\binom{m}{K}\binom{N-m}{n-K}}{\binom{N}{n}} \qquad E[X] = n \cdot \frac{m}{N} \qquad V(X) = n \cdot \frac{m}{N} \cdot \frac{(N-m)}{N} \frac{(N-n)}{N-1}$$

Poisson Random Variable: Fix any positive number $\lambda > 0$. Random variable X taking values $0, 1, 2, \ldots$ with probabilities $P(X=K) = e^{-\lambda} \cdot \frac{\lambda^K}{K!}$ is called a Poisson Random Variable with parameter λ.

Note that $\sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$

Theorem: Let X be a Poisson Random Variable with parameter λ.

$E[X] = \lambda$, $V(X) = \lambda$ } These will be given for all rand. vars. in the exam.

Proof: Using the formula $\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$

$\boxed{Ex}$ A manager wants to buy machine A or B. The daily repair required for A is X which is a poisson rand var with $\lambda = 0.1t$ where t is the time (in hours) of daily operations. The daily repair required for B is Y which is poisson rand var. with mean $0.12t$. The daily cost of operatings are

$$\left.\begin{aligned} C_A(t) &= 10t + 30X^2 \\ C_B(t) &= 8t + 30Y^2 \end{aligned}\right\}$$

Which machine costs less if, (a) a day is $t = 10$ hours (b) a day is $t = 20$ hours.

$(E[X])$

(a) $t=10 \Rightarrow C_A(10) = 100 + 30X^2$, $\lambda = 0.1t = 1$

$E[C_A] = 100 + 30\,E[X^2] = 100 + 30 \cdot 2 = \underline{\underline{160}}$

$V(X) = E[X^2] - E[X]^2 \Rightarrow 1 = E[X^2] - 1 \Rightarrow E[X^2] = 2$

$t=10 \Rightarrow C_B(10) = 80 + 30Y^2$, $\lambda = 0.12t = 1.2$

$E[C_B] = 80 + 30\,E[Y^2] = 80 + 30\left[(1.2)^2 + 1.2\right] = \underline{\underline{159.2}}$

$V(X) = E[X^2] - E[X]^2 \Rightarrow E[Y^2] = (1.2)^2 + 1.2$

(b) <same operations> ... $E[C_A] = \underline{\underline{380}}$ $E[C_B] = \underline{\underline{404.8}}$

So A is economic for $(t=20)$, B is economic for $(t=10)$.

Remark: The Poisson can be used as an approximation for a binomial random variable with parameters n and p when n is large, p is small and so np is moderate.

$\lambda = E[X] = np$ should be taken.

Ex) Suppose prob. of a misprint per page is 0.001. A book contains 1000 page. Find the probability that exactly 5 misprint pages.

$X = \#\text{misprints}$, $P(X=5) = \binom{1000}{5} \cdot 0.001^5 \cdot 0.999^{995} = ?$

$P(X=5) \approx e^{-\lambda} \cdot \frac{\lambda^5}{5!} = e^{-1} \cdot \frac{1^5}{5!} = \frac{1}{120e} =$

$P(Y=5)$ — Poisson

$\lambda = np = 1$

(*) Let us suppose that events are occuring at certain (random points) of time interval T and that for some positive constant λ the following assumptions hold true.

(1) The probability that exactly one event occurs in a given interval of time Δt is $\lambda \Delta t + o(\Delta t)$ when $\Delta t \to 0$

(2) The points that ≥ 2 events occur in time interval Δt is $o(\Delta t)$ when $\Delta t \to 0$.

(3) The probability of an occuring event during time interval does not depend on what happened prior to that time.

→ Then, the number of events occuring during a time interval of length t is approximately, a Poisson random variable with mean λt.

Ex) An average of 10 car/minute cars are passing.

(a) Find the probability that at least 15 cars pass in the next minute.

(b) Find the probability that at least 15 cars pass in the next two minutes.

(a) $N(1) = \#$ cars during 1 minute

$N(1) =$ Poisson with $10 \Rightarrow P(N(1) \geq 15) = \sum_{k=15}^{\infty} e^{-10} \frac{10^k}{k!}$

<also can be written> $= 1 - \sum_{k=0}^{14} e^{-10} \frac{10^k}{k!}$

(b) $N(2) =$ Poisson with $20 \Rightarrow P(N(2) \geq 15) = \sum_{k=15}^{\infty} e^{-20} \frac{20^k}{k!}$

$= 1 - \sum_{k=0}^{14} e^{-20} \frac{20^k}{k!}$

Continuous Random Variable

Def: A random variable X is said to be continuous if it takes on all values on some interval and if there exists a function $f(x)$, called the probability density function such that;

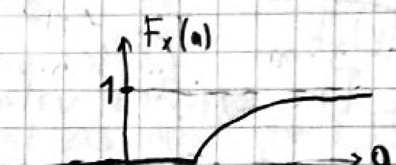
(i) $f(x) \geq 0 \ \forall$

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

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

Ex The prob. density function of X, the lifetime of a device (in hours) is given by the function $f(x) \begin{cases} \frac{10}{x^2} & x > 10 \\ 0 & x \leq 10 \end{cases}$. (a) Find $P(X > 20)$

(b) Find the distribution function $F_X(a) = P(X \leq a)$.



(a) $P(X > 20) = \int_{20}^{\infty} \frac{10}{x^2}\,dx = \left. \frac{-10}{x} \right]_{20}^{\infty} = \frac{1}{2}$

(b) $F_X(a) = P(X \leq a) = \begin{cases} 0 & a \leq 10 \\ \int_{10}^{a} f(x)\,dx & a > 10 \end{cases}$ $\quad \int_{10}^{a} \frac{10}{x^2}\,dx = \left. \frac{-10}{x} \right]_{10}^{a} = 1 - \frac{10}{a}$

Remark: If X is continuous random variable then, $P(X = a) = 0$

Also, $P(a \leq X \leq b) = P(a < X < b)$.

Ex The amount of hours that a computer functions before breaking is a cont. random variable with the prob density function given by,

$$f(x) = \begin{cases} \lambda e^{-\frac{x}{100}} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

(a) What is the prob of functioning 50-150 hours?

(b) " " " less than 100 hours?

$$1 = \int_{-\infty}^{\infty} f(x)dx \Rightarrow \int_0^{\infty} \lambda e^{-x/100} dx = \left[-100\lambda e^{-x/100}\right]_0^{\infty} = 100\lambda = 1 \Rightarrow \lambda = \frac{1}{100}$$

(a) $P(50 \leq X \leq 150) = \int_{50}^{150} \frac{1}{100} e^{-x/100} = \left[-e^{-x/100}\right]_{50}^{150} = -e^{-3/2} + e^{-1/2}$

(b) $P(X \leq 100) = \int_0^{100} \frac{1}{100} e^{-x/100} = \left[-e^{-x/100}\right]_0^{100} = -e^{-1} + 1 = 1 - 1/e$

Distribution function for cont. rand. vars: $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

$\Rightarrow$ $f(x) = F'(x)$ density can be calculated by derivative of dist. function.

Def: Let X be a continuous random variable with density function $f(x)$.

The expected value of X is, $E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$

The variance of X is, $V(X) = \int_{-\infty}^{\infty} (x - E[X])^2 f(x)\,dx$

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

The moment generating function for X is;

$$M(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$M'(t) = \frac{d}{dt}\int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_{-\infty}^{\infty} x e^{tx} f(x)dx \qquad M'(0) = E[X]$$

$$M''(t) = \frac{d^2}{dt^2}\int_{-\infty}^{+\infty} e^{tx} f(x)dx = \int_{-\infty}^{\infty} x^2 e^{tx} f(x)\,dx \qquad M''(0) = E[X^2]$$

$E[aX+b] = a \cdot E[X] + b$ $\qquad$ $V(aX+b) = a^2 \cdot V(X)$

$V(X) = E[X^2] - E[X]^2$ $\qquad$ as for the discrete variables.

Ex| The weekly demand X for a good at a certain supply station has a probability density function given by

$$f(x) = \begin{cases} x, & 0 \leq x \leq 1 \\ 1/2, & 1 < x \leq 2 \\ 0, & \text{otherwise} \end{cases}$$

Find $E[X]$ and $V(X)$ for X.

$$E[x] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x^2 dx + \int_1^2 \frac{x}{2} dx = \frac{x^3}{3}\Big]_0^1 + \frac{x^2}{4}\Big]_1^2 = \frac{1}{3} + 1 - \frac{1}{4} = \frac{13}{12}$$

$$V(X) = E[X^2] - E[X]^2 \Rightarrow E[x]^2 = \int_0^1 x^3 dx + \int_1^2 \frac{x^2}{2} dx = \frac{x^4}{4}\Big]_0^1 + \frac{x^3}{6}\Big]_1^2 = \frac{1}{4} + \frac{8}{6} - \frac{1}{6} = \frac{17}{12}$$

$$V(X) = \frac{17}{12} - \left(\frac{13}{12}\right)^2 = \frac{17 \cdot 12 - 13^2}{12^2} = \frac{35}{144}$$

Ex| Let X be a continuous rand var with density function,

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find $E[X]$ and $V[X]$

(b) Let $Y = 3X+1$, Find prob density $f_Y(y)$ for Y, $E[Y]$, $V[Y]$.

(a) $$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 2x^2 dx = \frac{2}{3}$$ $$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 2x^3 dx = \frac{1}{2}$$

$$V(X) = E[X^2] - E[X]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

(b) $$F_Y(y) = P(Y \leq y) = P(3X+1 \leq y) = P\left(X \leq \frac{y-1}{3}\right)$$

$$0 < x < 1 \Rightarrow 0 < \frac{y-1}{3} < 1 \quad F_Y(y) = \begin{cases} 0, & \text{if } y < 1 \\ 1, & \text{if } y > 4 \\ \int_{-\infty}^{\frac{y-1}{3}} f(x)\,dx = \int_0^{\frac{y-1}{3}} 2x\,dx = \left(\frac{y-1}{3}\right)^2 \end{cases}$$

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)\,dt \Rightarrow F_Y'(y) = f_Y(y)$$

$$\Rightarrow f_Y(y) = \begin{cases} \frac{2(y-1)}{9}, & 1 \leq y \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

SECOND SOLUTION

(b) $F_Y(y) = P\left(X \leq \frac{y-1}{3}\right) = F_X\left(\frac{y-1}{3}\right)$

$$f_Y(y) = \frac{d}{dy}\left(F_Y(y)\right) = \frac{d}{dy}\left(F_X\left(\frac{y-1}{3}\right)\right) = \langle \text{chain rule} \rangle =$$

$$= f_X\left(\frac{y-1}{3}\right)\cdot\frac{1}{3} = \begin{cases} 2\cdot\left(\frac{y-1}{3}\right)\cdot\frac{1}{3} & , \quad 0 < \frac{y-1}{3} < 1 \\ 0 & , \quad \text{otherwise} \end{cases} \quad //$$

$$E[Y] = E[3X+1] = 3E[X] + 1 = 3\cdot\frac{2}{3} + 1 = 3$$

$$V(Y) = V(3X+1) = 3^2\cdot V(X) = 9\cdot\frac{1}{18} = \frac{1}{2}$$

The Uniform Distribution

Def: A random variable X is said to be uniformly distributed over the integral over the interval (α, β) if its density function is

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & , \quad \alpha < x < \beta \\ 0 & , \quad \text{otherwise} \end{cases}$$

Theorem: For a uniformly distributed over (α, β) rand var X, we have

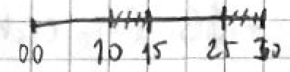
$$E[X] = \frac{\alpha+\beta}{2} \quad , \quad V(X) = \frac{(\alpha-\beta)^2}{12}$$

Ex: Buses arrive to a stop at 15 minutes interval starting 7:00 that is they arrive at 7:00, 7:15, 7:30... If a passenger arrives at the stop at a time uniformly distributed between 7:00 - 7:30. Find the prob of passenger wait,

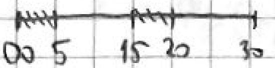
(a) less than 5 minutes.

(b) more than 10 minutes.

X = # minute past 7:00 the passenger arrives at the bus stop



$\Rightarrow$ (a) $P(10 < X < 15) + P(25 < X < 30) = \int_{10}^{15} \frac{1}{30}\,dx + \int_{25}^{30} \frac{1}{30}\,dx = \frac{1}{3}$ //



$\Rightarrow$ (b) $P(0 < X < 5) + P(15 < X < 20) = \frac{1}{3}$ //

Ex) Let X be uniformly distributed over interval (0, 10). Find $P(|X-5| > 4)$.

$\Rightarrow$ $P(0 < X < 1) + P(9 < X < 10) = \frac{1}{5}$ //

— Review —

① A fair die is tossed. If the number is even, it tossed 1 more time, if its odd, then a fair coin is tossed. Define $A = \{$a head appears on coin$\}$
$B = \{$the die is tossed one time$\}$

Find Ω, $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$, $P(A|B)$, $P(B|A)$, if A-B mutually exclusive.
if A-B independent events.

$$\Omega = \{(2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6), (1,H), (1,T), (3,H), (3,T), (5,H), (5,T)\}$$

$A = \{(1,H), (3,H), (5,H)\}$

$B = \{(1,H), (3,H), (5,H), (1,T), (3,T), (5,T)\}$

$P(A) = \frac{3}{6} \cdot \frac{1}{2} = \frac{1}{4}$ $\quad$ $P(B) = \frac{3}{6} = \frac{1}{2}$

$P(A \cap B) = P(A) = \frac{1}{4}$ $\quad$ $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{2}$

$P(A \cup B) = P(B) = \frac{1}{2}$ $\quad$ $P(B|A) = \frac{P(A \cap B)}{P(A)} = 1$

$P(A \cap B) \neq P(A)P(B) \Rightarrow$ not independent $\quad$ $A \cap B \neq \emptyset \Rightarrow$ not mutually exclusive

② Two balls are chosen randomly without replacement from 3 white, 4 yellow, 2 red balls. Suppose you get +\$4 for each white, −\$3 for each yellow ball selected. Let X be the amount we get.

(a) Find prob. dist. for X. (b) $E[X]=?$ (c) $V(x)=?$

3 white +4
4 yellow −3
2 red 0

(a)

	(YY)	(YR)	(RR)	(YW)	(WR)	(WW)
X	−6	−3	0	+1	4	8
P	$\frac{6}{36}$	$\frac{8}{36}$	$\frac{1}{36}$	$\frac{12}{36}$	$\frac{6}{36}$	$\frac{3}{36}$

$P(X=-6)=\dfrac{\binom{4}{2}}{\binom{9}{2}}=\dfrac{6}{36}$

$P(X=-3)=\dfrac{\binom{4}{1}\binom{2}{1}}{\binom{9}{2}}=\dfrac{8}{36}$

$P(X=0)=\dfrac{\binom{2}{2}}{\binom{9}{2}}=\dfrac{1}{36}$ $P(X=1)=\dfrac{\binom{3}{1}\binom{4}{1}}{\binom{9}{2}}=\dfrac{12}{36}$ $P(X=4)=\dfrac{\binom{3}{1}\binom{2}{1}}{\binom{9}{2}}=\dfrac{6}{36}$

$P(X=8)=\dfrac{\binom{3}{2}}{\binom{9}{2}}=\dfrac{3}{36}$

(b) $E[X]=-6\cdot\frac{1}{6}-3\cdot\frac{2}{9}+\frac{1}{3}+4\cdot\frac{1}{6}+8\cdot\frac{1}{12}=0$

(c) $V(X)=E[X^2]-E[X]^2=36\cdot\frac{1}{6}+9\cdot\frac{2}{9}+1\cdot\frac{1}{3}+16\cdot\frac{1}{6}+64\frac{1}{12}-0^2=\frac{49}{3}$

③ If 8 identical blackboards are divided among 4 schools,

(a) How many division are possible?

(b) How many divisions are possible, if each school gets at least 1.

(a) $x_1+x_2+x_3+x_4=8$ $\binom{r+n-1}{n-1}=\binom{8+4-1}{4-1}=\binom{11}{3}\longrightarrow$ Arrange $\underbrace{00000}_{8}\ \underbrace{||||}_{4}$

(b) $x_1+x_2+x_3+x_4=4$ $\binom{r+n-1}{n-1}=\binom{4+4-1}{4-1}=\binom{7}{3}=\frac{7!}{3!\,(7-3)!}$

④ An elevator starts at the basement with 8 people and discharges them by the time it reaches to the top floor 6.
and an operator

(a) In how many ways could the elevator operator have perceived the people leaving the elevator if all the people are alike to him?

(b) What if there are identical 5 men and 3 women as 8 people?

(a) $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 8 \Rightarrow \binom{r+n-1}{n-1} = \binom{8+6-1}{6-1} = \binom{13}{5} = \frac{13!}{5! \cdot 8!}$

(b) $\begin{matrix} x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 5 \\ y_1 + y_2 + y_3 + y_4 + y_5 + y_6 = 3 \end{matrix} \Rightarrow \binom{r+n-1}{n-1}_1 \cdot \binom{r+n-1}{n-1}_2 = \binom{10}{5} \cdot \binom{8}{5}$

5) We remove 12 face-cards (J, K, Q) from a deck so 40 cards remaining. We draw 5 cards with replacement.

(a) What is the prob. of getting at least 1 ace? $P(X \geq 1) = 1 - P(X=0) = 1 - \left(\frac{9}{10}\right)^5$

(b) E[X] of num. of aces? $E[X] = n \cdot p = 5 \cdot \frac{1}{10} = 0.5$

(c) St. deviation of num. of aces? $\sqrt{V(x)} = \sqrt{np(1-p)} = \sqrt{5 \cdot \frac{1}{10}\left(1 - \frac{1}{10}\right)} = \frac{3\sqrt{5}}{10}$

6) A purchaser buys components in lots of size 10. It checks 3 items for each lot and accepts if all 3 are non-defective. If 30% of the lots have 4 defective components and 70% have 1 defective component, find what proportion of lots does the purchaser reject.

X = # defectives in the selected 3 components, $P(X \neq 0) = ?$

$$P(X=0) = P(x=0 \cap \overset{30\%}{\text{Type 1}}) + P(x=0 \cap \overset{70\%}{\text{Type 2}})$$

$$= P(\text{Type 1}) \cdot P(X=0 \mid \text{Type 1}) + P(\text{Type 2}) \cdot P(x=0 \mid \text{Type 2})$$

$$= \underbrace{0.3 \cdot \frac{\binom{6}{3}}{\binom{10}{3}} + 0.7 \cdot \frac{\binom{9}{3}}{\binom{10}{3}}}_{K}$$

$$P(X \neq 0) = P(X \geq 1) = 1 - P(X=0) = 1 - K$$

⑦ Let X be a Poisson Random variable. Show that

$$P(X \text{ is even}) = \frac{1}{2}\left(1+e^{-2\lambda}\right) \quad \text{where} \quad E[X]=\lambda.$$

$$P(X=K) = \frac{e^{-\lambda}\cdot\lambda^{K}}{K!}, \quad P(X \text{ is even}) = P(X=0)+P(X=2)+P(X=4)+\dots$$

$$= \sum_{K=0}^{\infty} e^{-\lambda}\cdot\frac{\lambda^{2K}}{(2K)!} =$$

$$e^{\lambda} = \sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = 1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\dots$$

$$e^{-\lambda} = \sum_{i=0}^{\infty}\frac{(-\lambda)^i}{i!} = 1-\lambda+\frac{\lambda^2}{2}-\frac{\lambda^3}{3}+\dots \Rightarrow \frac{e^{\lambda}+e^{-\lambda}}{2} = \text{(even i's)}$$

$$\Rightarrow\; = e^{-\lambda}\cdot\frac{e^{\lambda}+e^{-\lambda}}{2} = \frac{1+e^{-2\lambda}}{2} \;//$$

Sug6 #1| Let X be a random variable with prob. density function

$$f(x) = \begin{cases} C(1-x^2), & -1<x<1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find constant C.

(b) Dist func. for X = ?, $2X-1$ = ?

(c) What are the E and V for X and $2X-1$?

(a) $\int_{-1}^{1} C(1-x^2) = 1 \Rightarrow C\cdot\left[x-\frac{x^3}{3}\right]_{-1}^{1} = C\left[2-\frac{1}{3}-\frac{1}{3}\right]=1 \Rightarrow C=\frac{3}{4}$ //

(b) $F_X(x)=0$ for $x<-1$

$F_X(x)=1$ for $x>1$

$$F_x(x) = P(X\leq x) = \int_{-1}^{x}\frac{3}{4}(1-t^2)\,dt = \underbrace{\frac{-1}{4}x^3+\frac{3}{4}x+\frac{1}{2}}_{\text{for } -1<x<1}$$

$F_Y(x)=0$ for $x<-$

$F_Y(x)=1$ for $x>1$

$$F_Y(x) = P(2X-1\leq x) =$$

$$X \leq \frac{x+1}{2}$$

(*) changing continuous rand. vars

$X \rightarrow 2X-1$

(c) $E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{-1}^{1} x \frac{3}{4}(1-x^2)\, dx = \int_{-1}^{1} \frac{3}{4}x - \frac{3}{4}x^3 = 0$

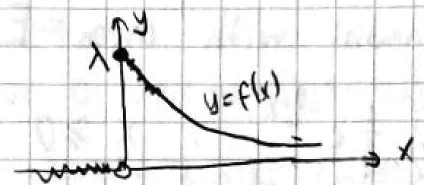
$V(X) = E[X^2] - \underbrace{E[X]^2}_{0} = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{-1}^{1} x^2 \frac{3}{4}(1-x^2)\, dx = \frac{1}{5}$

$E[2X-1] = 2E[X] - 1 = -1 \qquad V(2X-1) = 2^2 \cdot V(X) = \frac{4}{5}$

The Exponential Distribution

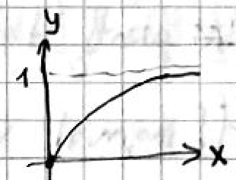
Def: A continuous rand. var X is called exponential with parameter $\lambda > 0$ if its density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



Cumulative dist. function of exponential;

$$F(x) = P(X \leq x) \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$



Theorem: For exponential random variable X with parameter λ;

$$E[X] = \frac{1}{\lambda}, \qquad V(X) = \frac{1}{\lambda^2}$$

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{0}^{\infty} x \cdot \lambda e^{-\lambda x}\, dx = \left(\begin{matrix}\text{integration} \\ \text{by parts}\end{matrix}\right) = \frac{1}{\lambda}$$

Ex| A sugar refinary has three processing plants. The amount of sugar that one plant can process in one day can be modeled as having exponential distribution with a parameter 4 (in tons.) If the plant's are operating independently, find the probability that exactly 2 of the three plants process more than 4 tons each on a given day.

X = The amount of sugar that one plant processes on the day.

X is exponential with mean $= E[X] = \frac{1}{\lambda} = 4 \Rightarrow \lambda = \frac{1}{4}$

$$f(x) = \begin{cases} \frac{1}{4} e^{-\frac{1}{4}x} & , \; x \geq 0 \\ 0 & , \; x < 0 \end{cases}$$ is its density function.

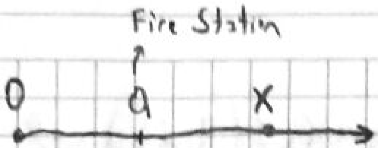
Y = # plants that process more than 4 tons for the day. $P(Y=2) = ?$

$= (\text{Binomial with } n=3, \; p = P(X>4))$

$$P(X>4) = \int_4^\infty \frac{1}{4} e^{-\frac{1}{4}x} = \left[\frac{\frac{1}{4}}{-\frac{1}{4}} e^{-\frac{1}{4}x} \right]_4^\infty = 0 + e^{-1} = \frac{1}{e}$$

$$P(Y=2) = \binom{n}{2} \cdot p^2 \cdot (1-p)^{n-2} = \binom{3}{2} \cdot \left(\frac{1}{e}\right)^2 \left(1 - \frac{1}{e}\right) = \frac{3(e-1)}{e^3} \; //$$

Ex| An infinite road from 0 to infinity. If the distance of a fire from 0 to ∞ is exponentially distributed with parameter λ, where should the fire station be placed.



Minimize $E[|X-a|]$ where X is exp. dist. with λ.

$$E[|X-a|] = \int_{-\infty}^{\infty} |x-a| \cdot \underbrace{f(x)}_{\text{density}} dx = \int_0^{\infty} |x-a| \cdot \lambda e^{-\lambda x} dx$$ } Find a value that minimizes this.

$$= \int_0^a (a-x)\lambda e^{-\lambda x} dx + \int_a^{\infty} (x-a)\lambda e^{-\lambda x} = \left(\begin{matrix}\text{integration}\\ \text{by part}\end{matrix}\right) = a + \frac{2}{\lambda}\left(e^{-\lambda a} - \frac{1}{2}\right) = h(a)$$

$$h'(a) = 1 - 2e^{-\lambda a} = 0 \Rightarrow a = \frac{\ln 2}{\lambda} \,// \qquad \{ h''(a) = 2\lambda e^{-\lambda a} > 0 \}$$

The Gamma Distribution

Def: A rand. var. X is said to have a gamma distribution with parameters (t, λ), $t > 0$, $\lambda > 0$ if its density function is,

$$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x} (\lambda x)^{t-1}}{\Gamma(t)}, & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad \Gamma(t) = \int_0^{\infty} e^{-y} y^{t-1} dy$$

⊛ Note that $$1 = \int_0^{\infty} \frac{\lambda e^{-\lambda x} (\lambda x)^{t-1}}{\Gamma(t)} dx \Longrightarrow \int_0^{\infty} e^{-\lambda x} x^{t-1} dx = \frac{\Gamma(t)}{\lambda^t}$$

⊛ Note that $\Gamma(1) = 1$, $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \Rightarrow$ density function for exponential

⊛ $\Gamma(t+1) = t \cdot \Gamma(t)$, In natural numbers, $\Gamma(n) = (n-1)!$

$\Gamma(0.5) = \sqrt{\pi}$

Theorem: Let X be gamma rand. var. with params (t, λ).

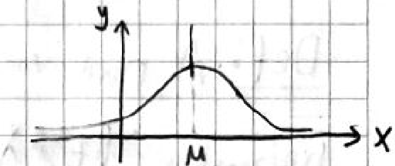
$$E[X] = \frac{t}{\lambda} \qquad V(X) = \frac{t}{\lambda^2}$$

$$E[X] = \int_0^{\infty} x \, \frac{\lambda e^{-\lambda x} (\lambda x)^{t-1}}{\Gamma(t)} \, dx = \frac{\Gamma(t+1)}{\Gamma(t)} \cdot \frac{1}{\lambda} \cdot \overbrace{\int_0^{\infty} \frac{\lambda e^{-x} (\lambda x)^{t}}{\Gamma(t+1)} \, dx}^{1 \to \text{gamma with } (t+1, \lambda)} = \frac{t}{\lambda}$$

The Normal Distribution

Def: A rand. var. X is said to be normal with μ and $\sigma > 0$ if the density function is given by;

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-(x-\mu)^2 / 2\sigma^2}$$



Theorem: If X is normal with μ and σ, then

$$E[X] = \mu \qquad V(X) = \sigma^2$$

Theorem: Let X be normal with parameters μ, $\sigma > 0$. Then;

$Y = \alpha X + \beta$ is also normal with $(\alpha\mu + \beta)$, $(|\alpha|\sigma)$.

Proof: $F_Y(y) = P(Y \leq y) = P(\alpha X + \beta \leq y) = P\left(X \leq \frac{y-\beta}{\alpha}\right) = F_X\left(\frac{y-\beta}{\alpha}\right)$

Assume $\alpha > 0$

then, $f_Y(y) = \frac{d}{dy}\left(F_Y(y)\right) = \frac{d}{dy}\left(F_X\left(\frac{y-\beta}{\alpha}\right)\right) = f_X\left(\frac{y-\beta}{\alpha}\right) \cdot \frac{1}{\alpha} = \dots = \left(\begin{array}{c}\text{density for normal with} \\ \alpha\mu+\beta, \\ \alpha\sigma\end{array}\right)$

chain rule

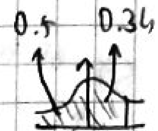
If $\alpha < 0 \Rightarrow$ we get $(-\alpha\sigma)$. $\Rightarrow \begin{bmatrix} \alpha\mu + \beta, \\ |\alpha|\sigma. \end{bmatrix}$

<u>Remark:</u> Let X be normal with μ and σ. Then, $Z = \frac{X-\mu}{\sigma}$ is also normal with parameters $E[Z]=0$, $V(Z)=1$. Such Z is called standard normal.

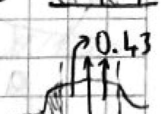
Note that $P(a \leq Z \leq b) = \int_a^b \frac{1}{2\pi} e^{\frac{-x^2}{2}} dx$ is impossible to evaluate. So we use the table.

Ex If Z is standard normal, find;

(a) $P(Z \leq 1) = 0.5 + P(0 \leq z \leq 1) = 0.5 + 0.3413 = 0.8413$



(b) $P(Z \leq -1.5) = 0.5 - P(0 \leq z \leq 1.5) = 0.5 - 0.4332 = 0.0668$



(c) $P(Z \geq 1.14) = 0.5 - P(0 \leq z \leq 1.14) = 0.5 - 0.3729 = 0.1271$

(d) $P(-1.56 \leq Z \leq 0.5) = P(0 \leq z \leq 1.56) + P(0 \leq z \leq 0.5) = 0.4406 + 0.1915 = 0.6321$

(★) If X is normal with μ, σ, then

standard normal

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right)$$

Ex A firm manufactures bottles and apple juice, has a machine that fills 16-ounce bottles automatically. If the ounces of fill per bottle is normally distributed with the average amount of 16 ounces and st. dev of 1 ounce. Find the prob. that machine will dispense more than 17 ounces in any bottle.

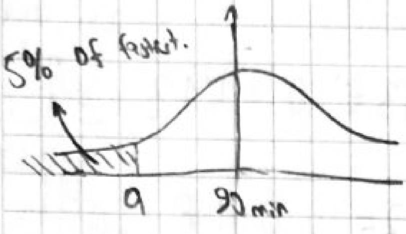
X = # ounces in a bottle $\Rightarrow$ normal with $\mu=16$, $\sigma=1$

$$P(X > 17) = P\left(\frac{X-\mu}{\sigma} > \frac{17-16}{1}\right) = P(Z > 1) = 0.5 - P(0 \leq z \leq 1) =$$

$$= 0.5 - 0.3413 = 0.1587$$

Ex An aptitude test for pilots requires a series of operation to be performed quickly. Suppose the time needed to complete the test is normally distributed with mean 90 min. and st. dev. 20 min. The top 5% of pilots get an honor certificate. How fast a pilot must complete the test to get the honor certificate.

X = # minutes to complete $\Rightarrow$ normal with $\mu = 90$, $\sigma = 20$.



We want to find a.

$$0.05 = P(X \leq a) = P\left(Z \leq \frac{a-90}{20}\right), \quad \text{Let } k = \frac{90-a}{20}$$

$$0.5 - P(0 \leq Z \leq k) = 0.05 \Rightarrow P(0 \leq Z \leq k) = 0.45 \approx P(0 \leq Z \leq 1.645)$$

$$k = \frac{90-a}{20} = 1.645 \Rightarrow \underline{a = 57.1 \text{ minutes}}$$

<u>Theorem</u> (Normal Prob. Rule): Let X be normally distributed with parameters μ and σ. Then;

$$P(\mu - \sigma < X < \mu + \sigma) \approx 0.68$$
$$P(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.95$$
$$P(\mu - 3\sigma < X < \mu + 3\sigma) \approx 0.997$$

<u>Proof:</u> $P(\mu - \sigma < X < \mu + \sigma) = P\left(-1 < \frac{X-\mu}{\sigma} < 1\right) = 2 \cdot P(0 < Z < 1) \approx 0.68$

$P(\mu - 2\sigma < X < \mu + 2\sigma) = \cdots$

$P(\mu - 3\sigma < X < \mu + 3\sigma) = \cdots$

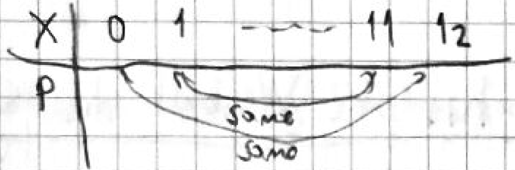
<u>Chebyshev's Theorem (remind):</u> For any random variable X, and $k>1$,

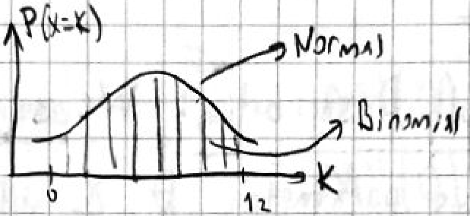
$$P(\mu - k\sigma < X < \mu - k\sigma) \geq 1 - \frac{1}{k^2}$$

<u>The Normal Approximation to Binomial Distribution:</u> Let X be Binomial with parameters n and p. Then,

$$P(X=k) = \binom{n}{k} p^k (1-p)^k \quad , \quad E[X] = np \, , \quad V(X) = np(1-p)$$

Let X be the number of tails we see in 12 coin flips.
Then, X is binomial with $p = \frac{1}{2}$ and $n = 12$.





⊛ If $np > 5$ and $n(1-p) > 5$ then, (binomial X) $\approx$ (normal Y)

$$P(X=k) = P\left(k - \frac{1}{2} < Y < k + \frac{1}{2}\right) \, , \quad E[Y] = np \, , \quad V(Y) = np(1-p)$$

<u>Ex</u> Find exact and approximate probability of getting 6 heads, 10 tails for 16 coin flips.

$$\binom{16}{6} \cdot \left(\frac{1}{2}\right)^6 \left(1 - \frac{1}{2}\right)^{10} = \underline{\underline{0.1222}} \quad , \quad \left.\begin{matrix} np = 8 > 5 \\ n(1-p) = 8 > 5 \end{matrix}\right\} \Rightarrow \text{normal } Y$$

$$P(X=6) \approx P\left(6 - \frac{1}{2} < Y < 6 + \frac{1}{2}\right) =$$

$$P(5.5 < Y < 6.5) = P\left(\frac{5.5-8}{2} < \frac{Y-\mu}{\sigma} < \frac{6.5-8}{2}\right)$$

$$= P(-1.25 < Z < -0.75)$$

$$= \underline{0.1210}$$

$$Y \left\{ \begin{matrix} \mu = np = 8 \\ \sigma^2 = np(1-p) \\ \sigma = 2 \end{matrix} \right.$$

Ex) What is the prob. that at least 26 of 50 mosquitoes will be killed by a new insect spray when the prob of a killing is 0.6.

X = # mosquitoes killed among 50. X is binomial with $p=0.6$, $n=50$.

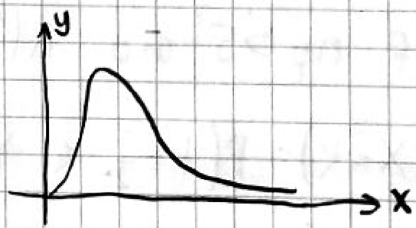
$$P(X \geq 26) = P(\overset{normal}{Y} > 25.5) \quad Y \text{ is normal with } \begin{array}{l} \mu = np = 30 \\ \sigma = \sqrt{np(1-p)} = 3.464 \end{array}$$

$$= P\left(\frac{Y-\mu}{\sigma} \geq \frac{25.5-30}{3.464}\right) = P(Z \geq -1.3) = 0.5 + P(0 \leq Z < 1.3) = 0.9032$$

* We can put $50 \geq X \geq 26$ but it won't change the result significantly.

The Weibull Distribution: We say X has the Weibull distribution with positive parameters θ, γ if its density is given by;

$$f(x) \begin{cases} \frac{\gamma}{\theta} x^{\gamma-1} e^{-\frac{x^\gamma}{\theta}} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$



Note that, $$\int_0^\infty \frac{\gamma}{\theta} x^{\gamma-1} e^{\frac{-x^\gamma}{\theta}} dx = \left[-e^{\frac{-x^\gamma}{\theta}}\right]_0^\infty = -(-e^0) = 1$$

Theorem: If X has Weibull dist with params θ, γ, then

$$E[X] = \theta^{\frac{1}{\gamma}} \cdot \Gamma\left(1+\frac{1}{\gamma}\right) \qquad V(X) = \theta^{\frac{2}{\gamma}}\left[\Gamma\left(1+\frac{2}{\gamma}\right) - \left(\Gamma\left(1+\frac{1}{\gamma}\right)\right)^2\right]$$

where $\Gamma(t) = \int_0^\infty e^{-y} y^{t-1} dy$ is the Gamma function.

Ex| The length of service time of a product has been observed to follow a Weibull dist. with $\theta=50$, $\gamma=2$.

(a) Find the prob. that one of this product will function over 10 hours?

(b) Find the expected life length of this product.

$X=$ total service time of a product, X is weibull with $\theta=50$, $\gamma=2$

(a) $P(X>10)=\int_{10}^{\infty}\frac{2x}{50}e^{-\frac{x^2}{50}}dx=\left[-e^{-\frac{x^2}{50}}\right]_{10}^{\infty}=-\left(-e^{-\frac{100}{50}}\right)=e^{-2}$

(b) $E[X]=\theta^{\frac{1}{\gamma}}\cdot\Gamma\left(1+\frac{1}{\gamma}\right)=50^{\frac{1}{2}}\cdot\Gamma\left(\frac{3}{2}\right)=\sqrt{50}\cdot\left(\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right)\right)=\frac{\sqrt{50\pi}}{2}$

Moment Generating Functions for Continuous:

$$M(t)=E[e^{tX}]=\int_{-\infty}^{\infty}e^{tx}f(x)\,dx$$

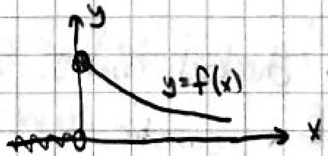
$M'(0)=E[X]$

$M''(0)=E[X^2]$

$M^{(n)}(0)=E[X^n]$

Ex| Find $M(t)$ for the exponential random variable with the density.

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$$



when t is close to 0.

$$M(t)=\int_0^{\infty}e^{tx}\cdot\lambda e^{-\lambda x}\,dx=\frac{\lambda}{t-\lambda}\cdot e^{x(t-\lambda)}\Big]_{x=0}^{x=\infty}=\frac{\lambda}{\lambda-t}$$

$M'(t)=\frac{\lambda}{(\lambda-t)^2}\Rightarrow E[X]=M'(0)=\frac{1}{\lambda}$

$M''(t)=\frac{2\lambda}{(\lambda-t)^3}\Longrightarrow E[X^2]=M''(0)=$ - -

<u>Joint and Marginal Prob. Dist.:</u> Let X_1, X_2 be discrete random variables. The joint prob. dist of X_1 and X_2 is given by,

$$p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$$

This function $p(x_1, x_2)$ is called the joint density of X_1 and X_2.

The marginal prob. function for X_1 is $p_1(x_1) = \sum_{x_2} p(x_1, x_2)$

The " " " X_2 is $p_2(x_2) = \sum_{x_1} p(x_1, x_2)$

X_2 \ X_1	a_1	a_2	---	a_n
b_1	$p(a_1,b_1)$	$p(a_2,b_1)$	...	$p(a_n, b_1)$
b_2				
b_n				$p(a_n, b_n)$

Marginal Probs for X_2:

$p_2(b_1) = p(a_1, b_1) + p(a_2, b_1) + \dots + p(a_n, b_1) = P(X_2 = b_1)$

$p_2(b_2) = p(a_1, b_2) + p(a_2, b_2) + \dots + p(a_n, b_2) = P(X_2 = b_2)$

Marginal probs for X_1: $p_1(a_1) = p(a_1, b_1) + p(a_1, b_2) + \dots + p(a_1, b_n) = P(X_1 = a_1)$

$\boxed{\text{Ex}}$ There are 3 checkout counters. 2 customers arrive at different times and select a random counter.

Let X_1 = # times Counter 1 is selected.
Let X_2 = # times Counter 2 is selected.

(a) Find the joint prob dist of X_1 and X_2

(b) Find the prob that one of the customers visits Counter2 given that one is known to visited C1.

X_2 \ X_1	0	1	2	Marg for X_2
0	1/9	2/9	1/9	4/9
1	2/9	2/9	0	4/9
2	1/9	0	0	1/9
Marg for X_1	4/9	4/9	1/9	

$p(0,0) = \frac{1}{9}$

$p(0,1) = \frac{2}{9}$

$p(7, \frac{1}{2}) = 0$

$p(-3, -3.14) = 0$

$p(x_1, x_2)$ is defined on whole cartesian coordinates.

Def: Let X_1, X_2 be continuous random variables. The joint prob. dist. of X_1 and X_2 is given by a nonnegative function $f(x_1, x_2)$ such that;

$$P(a < X_1 < b,\ c < X_2 < d) = \int_c^d \int_a^b f(x_1, x_2)\, dx_1\, dx_2$$

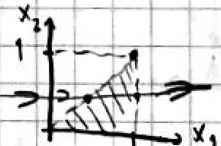
$$\left(\text{Note that } \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1\, dx_2 = 1\right)$$

The marginal prob. density function of X_1: $f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2$

The " " " X_2: $f_{X_2}(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1$

Ex Gasoline is stocked in a tank once each week and sold to customers. Let X_1 denote the proportion of the that is stocked in the tank for a particular week.
Let X_2 denote the proportion " is sold " "
($X_2 \leq X_1$)
It is known that the joint density of X_1, X_2 is given by,

$$f(x_1, x_2) = \begin{cases} 3x_1, & 0 \leq x_2 \leq x_1 \leq 1 \\ 0, & \text{otherwise} \end{cases}$$



Find the prob. that X_2 will be between 0.2 - 0.4 for a given week.

$$P(0.2 < X_2 < 0.4) = \int_{0.2}^{0.4} f_{X_2}(x_2)\, dx_2 = ?$$

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1 = \int_{x_2}^{1} 3x_1\, dx_1 = \left.\frac{3x_1^2}{2}\right]_{x_1=x_2}^{x_1=1} = \frac{3}{2}(1 - x_2^2)$$

$0 < x_2 < 1$

$$\Rightarrow f_{X_2}(x_2) = \begin{cases} \frac{3}{2}(1 - x_2^2), & 0 < x_2 < 1 \\ 0, & \text{otherwise} \end{cases} \Rightarrow \int_{0.2}^{0.4} \frac{3}{2}(1 - x_2^2)\, dx_2 = 0.272$$

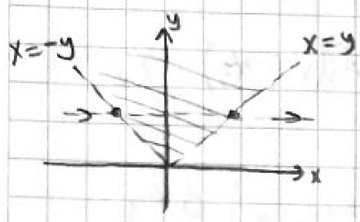
Ex | The joint prob. density function of X and Y is given by,

$$f(x,y)= \begin{cases} C\cdot(y^2-x^2)e^{-y}, & -y\leq x\leq y,\ 0\leq y<\infty \\ 0, & \text{otherwise} \end{cases}$$

(a) Find constant C.

(b) Find prob. density function of X.

(c) Find $E[X]$.

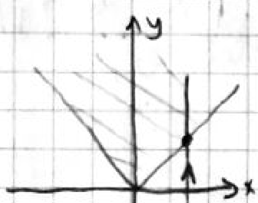
(a) 

$$1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy=\int_0^{\infty}\int_{-y}^{y} C\cdot(y^2-x^2)e^{-y}\,dx\,dy$$

$$=\int_0^{\infty} C\cdot\left(y^2e^{-y}x-\frac{x^3}{3}e^{-y}\right)\Bigg]_{x=-y}^{x=y} dy=\frac{4C}{3}\cdot\int_0^{\infty} y^3e^{-y}\,dy$$

$$=\frac{4}{3}C\cdot\Gamma(4)=\frac{4}{3}C\cdot 3!=8C=1 \implies C=\frac{1}{8}$$

(b) $$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy \underset{x>0}{=} \int_x^{\infty}\frac{1}{8}(y^2-x^2)e^{-y}\,dy=\frac{1}{4}e^{-x}(1+x)$$



$$\underset{x<0}{=}\int_{-x}^{\infty}\frac{1}{8}(y^2-x^2)e^{-y}\,dy=\frac{1}{4}e^{x}(1-x)$$

$$\implies f_X(x)=\frac{1}{4}e^{-|x|}(1+|x|)$$

(c) $$E[X]=\int_{-\infty}^{\infty} x\, f_X(x)\,dx=\int_{-\infty}^{\infty} x\cdot\frac{1}{4}e^{-|x|}(1+|x|)\,dx=0$$ Because the function is odd.

Def: (*) Discrete rand vars X_1, X_2 are independent if

$$P(X_1=x_1,\ X_2=x_2)=P(X_1=x_1)\cdot P(X_2=x_2) \quad \text{for all } x_1,\ x_2.$$

(**) Continuous rand vars X_1, X_2 are independent if,

$$f(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \quad \text{for all } x_1, x_2.$$

Ex | Let X_1 and X_2 be discrete with joint prob dst.

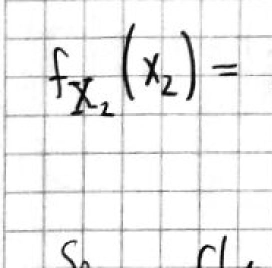
x_2 \ x_1	0	1	
0	0.38	0.17	0.55
1	0.14	0.02	0.16
2	0.24	0.05	0.29
	0.76	0.24	

$P(X_1=0, X_2=0) = 0.38 \neq 0.76 \cdot 0.55 = P(X_1=0) \cdot P(X_2=0)$

X_1 and X_2 are not independent.

Ex | Let X_1 and X_2 be continuous rand vars with joint prob density

$$f(x_1, x_2) = \begin{cases} 2(1-x_1) & , \ 0 \leq x_1 \leq 1, \ 0 \leq x_2 \leq 1 \\ 0 & , \ \text{otherwise.} \end{cases}$$



$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 = \underbrace{\int_0^1 2(1-x_1)\, dx_2}_{0<x_1<1} = \begin{cases} 2(1-x_1) & , \ 0<x_1<1 \\ 0 & , \ \text{otherwise} \end{cases}$$

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1 = \underbrace{\int_0^1 2(1-x_1)\, dx_1}_{0<x_2<1} = \begin{cases} 1 & , \ 0 \leq x_2 \leq 1 \\ 0 & , \ \text{otherwise} \end{cases}$$

Since $f(x_1, x_2) = 2(1-x_1) = [2(1-x_1)] \cdot [1] = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$

X_1 and X_2 are independent.

Quiz 4
1. joint of disc
2. joint of cont

Expected Values of Functions of Random Variables:

Def: Suppose the discrete rand. var.s X_1 and X_2 have a joint prob. function $p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$

The expected value of $g(X_1, X_2)$ is defined as

$$E[g(X_1, X_2)] = \sum_{x_2} \sum_{x_1} g(x_1, x_2) \cdot p(x_1, x_2)$$

Def: Suppose X_1 and X_2 are continuous random variables with joint density function $f(x_1, x_2)$. The expected value of $g(X_1, X_2)$ is

$$E[g(X_1, X_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, x_2) \cdot f(x_1, x_2)\, dx_1\, dx_2$$

Ex) $E[X_1 + X_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 + x_2)\, f(x_1, x_2)\, dx_1\, dx_2$

Theorem: If X_1 and X_2 are independent,

$$E[g(X_1) \cdot h(X_2)] = E[g(X_1)] \cdot E[h(X_2)]$$

Def: The covariance of X_1 and X_2 is defined as

$$Cov(X_1, X_2) = E[(X_1 - \mu_1) \cdot (X_2 - \mu_2)]$$

where $\mu_1 = E[X_1]$ and $\mu_2 = E[X_2]$

Remark: $Cov(X_1, X_1) = E[(X_1 - \mu_1)^2] = V(X_1)$

Theorem: $Cov(X_1, X_2) = E[X_1 \cdot X_2] - (E[X_1] \cdot E[X_2])$

Corollary: If X_1 and X_2 are independent, $Cov(X_1, X_2) = 0$.

Ex) Consider joint prob. dist. of X_1 and X_2 below.

	X_1: -1	0	1	Marg of X_2
X_2: -1	$\frac{1}{8}$	$1/8$	$1/8$	$3/8$
0	$\frac{1}{8}$	0	$1/8$	$2/8$
1	$\frac{1}{8}$	$1/8$	$1/8$	$3/8$
Marg fr X_2	$3/8$	$2/8$	$3/8$	1

(a) Are X_1, X_2 independent?

$$P(X_1=-1, X_2=-1) \neq P(X_1=-1) \cdot P(X_2=-1)$$
$$(1/8) \neq (3/8) \cdot (3/8)$$

(b) $Cov(X_1, X_2) = ?$

$$= E[X_1 \cdot X_2] - (E[X_1] \cdot E[X_2]) = ?$$

$$E[X_1 X_2] = (-1)(-1) \cdot \frac{1}{8} + 0 \cdot (-1) \cdot \frac{1}{8} + 1(-1) \cdot \frac{1}{8} +$$
$$(-1) \cdot 0 \cdot \frac{1}{8} + 0 \cdot 0 \cdot 0 + 1 \cdot 0 \cdot \frac{1}{8} +$$
$$(-1) \cdot 1 \cdot \frac{1}{8} + 0 \cdot 1 \cdot \frac{1}{8} + 1 \cdot 1 \cdot \frac{1}{8} = \frac{1}{8} - \frac{1}{8} - \frac{1}{8} + \frac{1}{8} = 0$$

$$E[X_1] = (-1)\frac{3}{8} + 0 \cdot \frac{2}{8} + 1 \cdot \frac{3}{8} = 0$$
$$E[X_2] = (-1)\frac{3}{8} + 0 \cdot \frac{2}{8} + 1 \cdot \frac{3}{8} = 0$$

$\Rightarrow Cov(X_1, X_2) = 0$

but; X_1, X_2 are not independent.

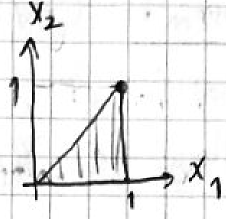
Review

Ex) With X_1 denoting the number of gasoline stocked at the beginning of a week, X_2 denoting the amount sold during that week.

$Y = X_1 - X_2$, represents the gasoline left.

The joint density function of X_1, X_2 is given by

$$f(x_1, x_2)\begin{cases} 3x_1 , & 0 \le x_2 \le x_1 \le 1 \\ 0 , & \text{otherwise} \end{cases}$$



Find $E[Y]$ and $V(Y)$ values.

(a) Method 1: $E[Y] = E[X_1 - X_2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - x_2)\, f(x_1, x_2)\, dx_1\, dx_2 = \dots$

Method 2: $E[Y] = E[X_1 - X_2] = E[X_1] - E[X_2] = \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)\, dx_1 - \int_{-\infty}^{\infty} x_2 f_{X_2}(x_2)\, dx_2$

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 \underset{0<x_1<1}{=} \int_0^{x_1} 3x_1\, dx_2 = 3x_1 x_2 \Big]_{x_2=0}^{x_2=x_1} = \begin{cases} 3x_1^2 & \text{for } 0<x_1<1 \\ 0 & \text{for otherwise.} \end{cases}$$

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1 \underset{0<x_2<1}{=} \int_{x_2}^{1} 3x_1\, dx_1 = \frac{3x_1^2}{2}\Big]_{x_1=x_2}^{x_1=1} = \begin{cases} \frac{3}{2}(1-x_2^2) & \text{for } 0<x_2<1 \\ 0 & \text{for otherwise.} \end{cases}$$

$$E[X_1] = \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)\, dx_1 = \int_0^1 3x_1^3\, dx_1 = \frac{3x_1^4}{4}\Big]_0^1 = \frac{3}{4}$$

$$E[X_2] = \int_{-\infty}^{\infty} x_2 f_{X_2}(x_2)\, dx_2 = \int_0^1 \frac{3}{2}(x_2^1 - x_2^3)\, dx_2 = \frac{3}{2}\left(\frac{x_2^2}{2} - \frac{x_2^4}{4}\right)\Big]_0^1 = \frac{3}{8}$$

$$E[Y] = \frac{3}{4} - \frac{3}{8} = \frac{3}{8} //$$

(b) $V(Y) = E[(Y-\mu)^2] = E[(X_1 - X_2 - (\mu_1 - \mu_2))^2] = E[((X_1-\mu_1) - (X_2-\mu_2))^2]$

$= E[(X_1-\mu_1)^2] + E[(X_2-\mu_2)^2] + E[(x_1-\mu_1)(x_2-\mu_2)] = V(X_1) + V(X_2) - 2\,Cov(X_1, X_2)$

$$V(X_1) = E[X_1^2] - (E[X_1])^2 = \int_0^1 x_1^2 \cdot 3x_1^2\, dx_1 - \left(\frac{3}{4}\right)^2 = \frac{3}{5} - \frac{9}{16}$$

$$V(X_2) = E[X_2^2] - (E[X_2])^2 = \int_0^1 x_2^2\left(\frac{3}{2} - \frac{3}{2}x_2^2\right) dx_2 - \left(\frac{3}{8}\right)^2 = \frac{1}{5} - \frac{9}{64}$$

$$Cov(X_1, X_2) = E[X_1X_2] - (E[X_1]E[X_2]) = \int_0^1\int_{x_2}^1 x_1 x_2 (3x_1)\, dx_1\, dx_2 - \frac{3}{4}\cdot\frac{3}{8} = \frac{3}{10} - \frac{9}{32}$$

select 3: 4 CS, 3 EE, 2 IE (9)

Ex (Old Exam) From a comittee consisting of 4 CS profs, 3 EE profs, and 2 IE profs, a subcomittee is randomly selected. Let,

$X_1 = \#$CS profs, $X_2 = \#$EE profs.

(a) Find joint prob dist of X_1 and X_2.

(b) Find the marginal dists of X_1 and X_2.

(c) Find the probability $P(X_1 = 1 \mid X_2 \geq 1)$.

(d) Find the value $Cov(X_1, X_2)$.

(a)(b)

X_2 \ X_1	0	1	2	3	marg of X_2
0	0	4/84	12/84	4/84	20/84
1	3/84	24/84	18/84	0	45/84
2	6/84	12/84	0	0	18/84
3	1/84	0	0	0	1/84
marg of X_1	10/84	40/84	30/84	4/84	1

$|\Omega| = \binom{9}{3} = 84$

$P(X_1 = 0, X_2 = 0) = 0$

$P(X_1 = 1, X_2 = 0) = \dfrac{\binom{4}{1}\binom{2}{2}}{84}$

$P(X_1 = 2, X_2 = 0) = \dfrac{\binom{4}{2}\cdot\binom{2}{1}}{84}$

$\vdots$

(c) $$P(X_1 = 1 \mid X_2 \geq 1) = \frac{P(X_1 = 1,\ X_2 \geq 1)}{P(X_2 \geq 1)} = \frac{\frac{24}{84} + \frac{12}{84}}{\frac{45}{84} + \frac{18}{84} + \frac{1}{84}} = \frac{36}{64} = \frac{9}{16}$$

(d) $$E[X_1 X_2] = 1\cdot 1\cdot\frac{24}{84} + 1\cdot 2\cdot\frac{18}{84} + 2\cdot 1\cdot\frac{12}{84} = \frac{84}{84} = 1$$

$$E[X_1] = 1\cdot\frac{40}{84} + 2\cdot\frac{30}{84} + 3\cdot\frac{4}{84} = \frac{112}{84} = \frac{4}{3}$$

$$E[X_2] = 1\cdot\frac{45}{84} + 2\cdot\frac{18}{84} + 3\cdot\frac{1}{84} = \frac{84}{84} = 1$$

$$Cov(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2] = 1 - \frac{4}{3}\cdot 1 = -\frac{1}{3}$$

Ex | The time it takes for a student to answer all questions on an exam is an exponential random variable with mean 1 hour 15 minutes. If we take 10 students, what is the prob. at least 1 of them completes the exam less than 1 hour.

$Y = $ #students among 10, completing <1 hour. $P(Y \geq 1) = ?$

$=$ Binomial with $n = 10$, $p = P(X < 60)$ where X is defined as

$X = $ time to complete the exam $=$ Exponential with $\beta = 75$ $(\lambda = 1/75)$

$$P(X<60) = \int_0^{60} \frac{1}{75} \cdot e^{-\frac{1}{75}x} dx = -e^{-\frac{1}{75}x}\Big]_0^{60} = 1 - e^{-\frac{60}{75}} = 1 - e^{-\frac{4}{5}}$$

$$P(Y \geq 1) = 1 - P(Y=0) = 1 - \left(1 - (1 - e^{-4/5})\right)^{10} = 1 - e^{-8}$$

Ex | Your travel time from home to office is normally distributed with mean 40 mins and st. dev 7 min. If you want to be 95% certain that you will not be late for an office appointment at 13:00 what is the latest time you should leave home?

$X = $ Travel time $=$ Normal with $\mu = 40$, $\sigma = 7$

$$P(X < k) = 95\% \Rightarrow P\left(\frac{X-\mu}{\sigma} < \frac{k-\mu}{\sigma}\right) = P\left(Z < \frac{k-40}{7}\right) = 0.95$$

Time it takes to travel

$$= 0.5 + P\left(0 < Z < \frac{k-40}{7}\right) = 0.95 \quad P\left(0 < Z < \frac{k-40}{7}\right) = 0.45$$

$$\Rightarrow \frac{k-40}{7} = 1.645 \Rightarrow k = 40 + 7 \cdot 1.645 = 52 \text{ mins} \Rightarrow 13:00 - 00:52 = \underline{\underline{12:08}}$$

Ex| Let X be standard normal ($\mu = 0$, $\sigma = 1$).

(a) Calculate $M(t)$. (b) Calculate $E[X^3]$, $E[X^4]$.

(a) $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$, $M(t) = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)\, dx =$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{tx - \frac{x^2}{2}} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-t)^2}{2}} \cdot e^{\frac{t^2}{2}}\, dx = e^{\frac{t^2}{2}} \cdot \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-t)^2}{2}}\, dx}_{\text{Normal with } \mu = t,\ \sigma = 1}$$

$$= e^{\frac{t^2}{2}} \cdot 1 \implies M(t) = e^{t^2/2}$$

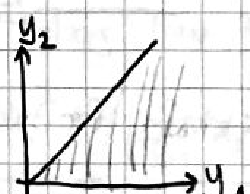
(b) $M'(t) = t \cdot e^{t^2/2}$ $\qquad M'''(t) = t \cdot e^{t^2/2} + 2t \cdot e^{t^2/2} + t^3 \cdot e^{t^2/2}$

$M''(t) = e^{t^2/2} + t^2 e^{t^2/2}$ $\qquad M''''(t) = 3e^{t^2/2} + 3t^2 e^{t^2/2} + 3t^2 e^{t^2/2} + t^4 e^{t^2/2}$

$\Rightarrow E[X^3] = 0$, $\quad E[X^4] = 3$

Ex| Let Y_1 and Y_2 have joint prob density function

$$f(y_1, y_2) = \begin{cases} e^{-y_1}, & 0 \leq y_2 \leq y_1 \leq \infty \\ 0, & \text{otherwise} \end{cases}$$

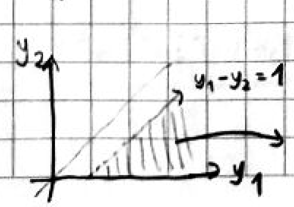


(a) $P(Y_1 - Y_2 \geq 1) = ?$

(b) Find marginal probability density functions for Y_1 and Y_2.

(c) Find $E[Y_1 - Y_2]$ and $V(Y_1 - Y_2)$.

(a) $P(Y_1 - Y_2 \geq 1) = \iint_{y_1 - y_2 \geq 1} f(y_1, y_2)\, dy_1\, dy_2 = \int_0^{\infty} \int_{1+y_2}^{\infty} e^{-y_1}\, dy_1\, dy_2$

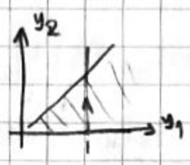


$y_1 - y_2 \geq 1$

$$= \int_0^{\infty} e^{-(y_2+1)}\, dy_2 = e^{-1}$$

(b) $f_{Y_1}(y_1)=\int_{-\infty}^{\infty} f(y_1,y_2)\,dy_2 \overset{y_1\geq 0}{=} \int_0^{y_1} e^{-y_1}\,dy_2 = y_1 e^{-y_1}$

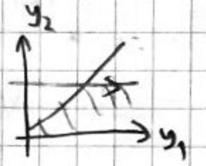
$$f_{Y_1}(y_1)=\begin{cases} y_1 e^{-y_1}, & y_1\geq 0\\ 0, & \text{otherwise}\end{cases}$$



Gamma with $\alpha=2$, $\beta=1 \Rightarrow E[Y_1]=\alpha\cdot\beta=2$

$f_{Y_2}(y_2)=\int_{-\infty}^{\infty} f(y_1,y_2)\,dy_1 \overset{y_2\geq 0}{=} \int_{y_2}^{\infty} e^{-y_1}\,dy_1 = e^{-y_2}$

$$f_{Y_2}(y_2)=\begin{cases} e^{-y_2}, & y_2\geq 0\\ 0, & \text{otherwise}\end{cases}$$



Exponential with $\beta=1 \Rightarrow E[Y_2]=\beta=1$

(c) $E[Y_1-Y_2]=E[Y_1]-E[Y_2]=2-1=1$

$V(Y_1-Y_2)=V(Y_1)+V(Y_2)-2\,Cov(Y_1,Y_2)$ ⊛ Exercise

——— *End of Review* ———

\# classes for a set with N data: $\lfloor \log_2 N \rfloor + 1$

$16 < N \leq 31$ 5
$32 \leq N \leq 63$ 6

⊛ Stem-Leaf display, frequency distribution, ogive

Random Sampling

Def: A random sample of size n from the distribution of X is a collection of n independent random variables $X_1, X_2, \dots, X_n$ each with same distribution as X

Def: Let $X_1, X_2, \dots, X_n$ be a random sample from the distribution X.

The statistic $\overline{X} = \dfrac{X_1 + X_2 + \dots + X_n}{n}$ is called a sample mean.

Ex) A random sample of 8 bulbs have life spans 866, 868, 848, 850, 851, 853, 842, 489.

$$\overline{X} = \frac{866+868+\dots+489}{8} = 808.375 \approx 808.4$$

Remark: We use ONE more decimal place for the mean.

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

The statistic
$$S^2 = \frac{(X_1-\overline{X})^2 + (X_2-\overline{X})^2 + \dots + (X_n-\overline{X})^2}{n-1}$$

is called sample variance. The sample standard dev is $S = \sqrt{S^2}$.

Remark: Variance is reported with TWO more decimal places from the data numbers.

(*) Computational formula for sample variance:
$$\boxed{S^2 = \frac{\sum_{k=1}^{n} X_k^2}{n-1} - \frac{\left(\sum_{k=1}^{n} X_k\right)^2}{n(n-1)}}$$

Ex) For the sample consisting of observations 2, 3, 4, 7, 9

$$\overline{X} = \frac{2+3+4+7+9}{5} = 5 \qquad S^2 = \frac{(2-5)^2 + (3-5)^2 + (4-5)^2 + (7-5)^2 + (9-5)^2}{5-1} = 8.5$$

$$S^2 = \frac{2^2+3^2+4^2+7^2+9^2}{5-1} - \frac{(2+3+4+7+9)^2}{5(5-1)} = 8.5$$

Finding the Sample Interquartile Range:

1. Find the median location $\left(\frac{n+1}{2}\right)$.

Median is denoted with $\tilde{x}$.
$1, 2 \} 3, 4 \Rightarrow \tilde{x} = 2.5$
$1, 3, 4 \Rightarrow \tilde{x} = 3$

2. Truncate the median location by rounding down.
3. Find the quartile location. $q = \frac{\text{Truncated Median} + 1}{2}$
4. Find q_1 by counting up from the smallest location to q.
   Find q_3 by counting down from the largest data to q.
5. Define interquartile range by $iqr = q_3 - q_1$

Ex| For sample 1, 2, 3, 4, 5 $trm = \frac{5+1}{2} = 3$ $q = \frac{3+1}{2} = 2$ $\to q_1 = 2$, $\to q_3 = 4$ $iqr = 2$

Ex| For sample 1,2,3,4,5,6,7,8 $trm = \left\lfloor \frac{8+1}{2} \right\rfloor = 4$ $q = \frac{4+1}{2} = 2.5$ $\to q_1 = 2.5$, $\to q_3 = 6.5$ $iqr = 4$

Constructing Box Plot:

2. Find $\tilde{x}$, q_1, q_2, iqr
3. Find inner fences f_1 and f_3. $f_1 = q_1 - 1.5(iqr)$ $f_3 = q_3 + 1.5(iqr)$
4. Find two points a_1, a_3 (adjacent values).
   a_1 is the data point closest to f_1 but $f_1 \leq a_1$
   a_3 is the data point closest to f_3 but $a_3 \leq f_3$
5. Find outer fences F_1 and F_3.

$$F_1 = q_1 - 2 * 1.5(iqr)$$
$$F_3 = q_3 + 2 * 1.5(iqr)$$

Ex For the data set given in stem-leaf display, construct the boxplot.

0	8
1	2
2	0 7
3	0 2 5 6
4	0 0 0 0 1 2 5 7
5	0 2
6	1
7	
8	9
9	
10	8

Dataset is: 8, 12, 20, 27, 30, 32, 35, 36....

$Trm = \frac{21+1}{2} = 11$ $\tilde{x} = 40$ $q_1 = 32$, $q_2 = 47$

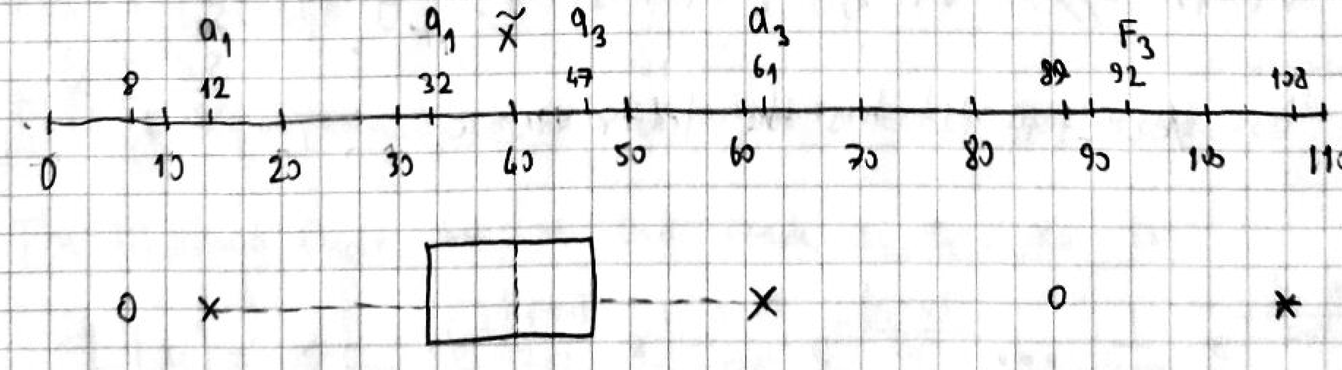
$q = \frac{11+1}{2} = 6$ $iqr = 47 - 32 = 15$

$f_1 = 32 - 1.5 \cdot 15 = 9.5 \Rightarrow a_1 = 12$

$f_3 = 47 + 1.5 * 15 = 69.5 \Rightarrow a_3 = 61$

$F_1 = 32 - 2 * 1.5 * 15 = -13$, $F_3 = 47 + 2 * 1.5 * 15 = 92$

6. Locate the points F_1, f_1, a_1, q_1, $\tilde{x}$, q_3, a_3, f_3, F_3 on the scale.
7. Construct the box with ends at q_1 and q_3 with an internal line drawn at the median.
8. Indicate adjacent values by x and connect them to the box.
9. Locate any data points between inner and outer fences and denote them by open circles. These points are considered to be mild outliers.
10. Indicate data points that fall beyond F_1, F_3 by asterix.



Point Estimators

Def: 1. A numerical feature of a population is called a parameter.
2. A statistic is a numerical valued function of the sample observation.
3. A statistic intended to estimate a parameter is called a point estimator.

(*) Let θ denote an arbitrary population parameter (such as $\mu, \sigma^2 \ldots$)
Let $\hat{\theta}$ denote an estimator of θ.

Def: 1. An estimator $\hat{\theta}$ is called unbiassed if $E[\hat{\theta}] = \theta$

2. If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiassed estimators of θ. Then the one with smaller variance is a more precis estimator.

Ex/ Let X_1, X_2, X_3, X_4, X_5 denotes a random sample from some population with $E[X_j] = \mu$ and $V(X_j) = \sigma^2$ for $j = 1 \ldots 5$. The following are suggested estimators for μ. $\hat{M}_1 = X_1$, $\hat{M}_2 = \frac{X_1 + X_2}{2}$ $\hat{M}_3 = \frac{X_1 + 2X_3}{2}$ $\hat{M}_4 = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$. Choose the best estimator.

biased

$E[\hat{M}_1] = E[X_1] = \mu$, $E[\hat{M}_2] = E\left[\frac{X_1 + X_2}{2}\right] = \mu$ $\cancel{E[\hat{M}_3] = \frac{3}{2}\mu}$ $E[M_4] = \mu$

$V(\hat{M}_1) = \sigma$ // , $V(\hat{M}_2) = \frac{1}{4} V(X_1 + X_2) = \frac{\sigma^2}{2}$ // $V(M_4) = \frac{1}{25} \cdot V(X_1 + X_2 + X_3 + X_4 + X_5)$

$= \frac{5\sigma^2}{25} = \frac{\sigma^2}{5}$ //

Use M_4 since it has the least variance.

The Method of Maximum Likelihood

1 question in final

Def: Suppose that X is a random variable with probability function $f(x,\theta)$ where θ is unknown. Let $x_1, x_2, \dots, x_n$ be the observed values in a random sample of size n. The likelihood function of θ of this sample is $L(\theta) = f(x_1,\theta)\cdot f(x_2,\theta)\cdot \dots \cdot f(x_n,\theta)$

The maximum likelihood estimator of θ is the value that maximizes likelihood function $L(\theta)$.

Ex) A box contains 4 balls, of which an unknown number θ are white. We are to sample two balls at random and count X, the number of white balls in the sample. Now, suppose we observe $X=1$. What value of θ will maximize the probability of this event?

θ white
$4-\theta$ black

$P(X=1 \mid \theta=0) = 0$
$P(X=1 \mid \theta=1) = 3/6$
$P(X=1 \mid \theta=2) = 4/6$ → Max prob ⟹ The maximum likelihood estimator of θ is 2.
$P(X=1 \mid \theta=3) = 3/6$
$P(X=1 \mid \theta=4) = 0$

Ex) Let X be normally distributed with unknown μ, known σ^2 (variance).

The likelihood function for μ and sample $x_1, x_2, \dots, x_n$ is

$$L(\mu) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x_2-\mu)^2}{2\sigma^2}} \cdots \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}$$

$$= L(\mu) = \frac{1}{\sigma^n \cdot (2\pi)^{n/2}} \cdot e^{-\frac{1}{2\sigma^2} \cdot \sum_{k=1}^{n} (x_k - \mu)^2}$$

$L(\mu)$ takes max value $\Updownarrow$ $\sum_{k=1}^{n} (x_k-\mu)^2$ takes min value

$$\Rightarrow \frac{d}{d\mu}\left(\sum_{k=1}^{n} (\mu - x_k)^2\right) = 0 \Rightarrow \sum_{k=1}^{n} 2(\mu - x_k) = 0$$

$$\Rightarrow 2\mu n - 2\sum_{k=1}^{n} x_k = 0 \Rightarrow \mu = \frac{\sum_{k=1}^{n} x_k}{n} \quad , \quad \underbrace{\frac{d^2}{d\mu^2}\left(\sum_{k=1}^{n} (\mu - x_k)^2\right) = 2n > 0}_{\text{The justification of critical point is unnecessary for this course.}}$$

$\mu_{critical}$

local min.

$$\hat{\mu} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Ex | Suppose we are to observe n independent lifelength measurements $X_1, X_2, \dots, X_n$ from components known to have lifelengths exhibiting a Weibull model given by

$$f(x) = \begin{cases} \frac{\gamma x^{\gamma - 1}}{\theta} e^{-\frac{x^\gamma}{\theta}} & , x > 0 \\ 0 & , x < 0 \end{cases}$$

Assume γ is known find the maximum likelihood estimator of θ.

$$L(\theta) = \frac{\gamma x_1^{\gamma-1}}{\theta} e^{-\frac{x_1^\gamma}{\theta}} \cdots \frac{\gamma x_n^{\gamma-1}}{\theta} e^{-\frac{x_n^\gamma}{\theta}} = \frac{\gamma^n x_1^{\gamma-1} x_2^{\gamma-1} \cdots x_n^{\gamma-1}}{\theta^n} \cdot e^{-\frac{1}{\theta}(x_1^\gamma + \dots + x_n^\gamma)}$$

$$\ln(L(\theta)) = \ln(\gamma^n x_1^{\gamma-1} \cdots x_n^{\gamma-1}) - n\ln(\theta) - \frac{1}{\theta}(x_1^\gamma + \dots + x_n^\gamma)$$

$$\frac{d}{d\theta}(\ln(L(\theta))) = -\frac{n}{\theta} + \frac{1}{\theta^2}(x_1^\gamma + \dots + x_n^\gamma) = 0 \iff \hat{\theta}_{mas} = \frac{X_1^\gamma + X_2^\gamma + \dots + X_n^\gamma}{n}$$

Check if its biased: $E[\hat{\theta}] = \frac{1}{n}(E[X_1] + \dots + E[X_n]) = \int_0^\infty x^\gamma \cdot \frac{\gamma - 1}{\theta} \cdot e^{-\frac{x^\gamma}{\theta}} dx = \dots$

$= \Gamma(2) \cdot \theta = \theta$

$\Rightarrow$ Then $\hat{\theta}$ is unbiassed.

Ex/ Let X be normally distributed with mean μ and var σ^2 where both μ, σ^2 are unknown. The likelihood function for μ and σ^2 for sample $x_1, x_2, \dots, x_n$ is

$$L(\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdots \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}} = \frac{e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2}}{(2\pi)^{n/2}\cdot\sigma^n}$$

In order to maximize L, we need to take partial derivatives.

$$\ln(L(\mu,\sigma^2)) = -\frac{1}{2\sigma^2}\sum_{k=1}^{n}(\mu-x_k)^2 - \frac{n}{2}\ln(2\pi) - n\ln(\sigma)$$

$$\frac{\partial \ln L(\mu,\sigma^2)}{\partial \mu} = -\frac{1}{2\sigma^2}\sum_{k=1}^{n}(\mu-x_k) = 0 \iff \underline{\underline{\mu_{max} = \frac{X_1+X_2+\dots+X_n}{n}}}$$

$$\frac{\partial \ln L(\mu,\sigma^2)}{\partial \sigma^2} = \frac{1}{2\sigma^4}\sum_{k=1}^{n}(\mu-x_k)^2 - \frac{n}{2}\cdot\frac{1}{\sigma^2}$$

$$= \frac{n}{2\sigma^2}\left(\frac{\sum_{k=1}^{n}(\mu-x_k)^2}{n} - \sigma^2\right) = 0 \iff \underline{\underline{\sigma^2_{max} = \frac{\sum_{k=1}^{n}(x_k-\mu)^2}{n}}}$$

Then we have $\hat{\mu} = \frac{X_1+X_2+\dots+X_n}{n}$, $\hat{\sigma^2} = \frac{\sum_{k=1}^{n}(x_k-\mu)^2}{n}$

It's possible to see $\hat{\mu}$ is unbiased and $\hat{\sigma^2}$ is biassed.

Interval Estimation and the Central Limit Theorem

Let $\overline{X}$ be a sample mean. We have the following:

$$E[\bar{X}] = E\left[\frac{X_1+X_2+\dots+X_n}{n}\right] = \frac{1}{n}\left(E[X_1]+\dots+E[X_n]\right) = \frac{1}{n}\cdot(\mu+\dots+\mu) = \mu$$

$$V(\bar{X}) = V\left(\frac{X_1+X_2+\dots+X_n}{n}\right) = \frac{1}{n^2}\left(V(X_1)+\dots+V(X_n)\right) = \frac{1}{n}\left(\sigma^2+\dots+\sigma^2\right) = \frac{\sigma^2}{n}$$

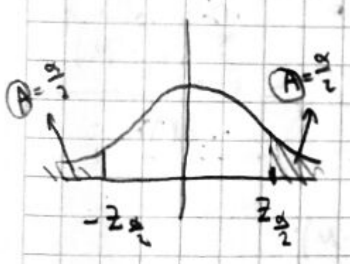
↳ $X_1, X_2, \dots X_n$ are independent.

Theorem: Let $X_1, X_2, \dots, X_n$ be a random sample from normal population with mean μ and variance σ^2. Then,

$\bar{X} = \dfrac{X_1+X_2+\dots+X_n}{n}$ is a normal rand. var. with $E[X]=\mu$, $V(X)=\dfrac{\sigma^2}{n}$

Proof: $M_{\bar{x}}(t) = E[e^{t\bar{x}}] = \dots$ = moment generating for normal with ↗. (long)

⊛ Denote by $Z_{\frac{\alpha}{2}}$ a positive number such that $P(Z \geq Z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$



$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = 1-\alpha$$

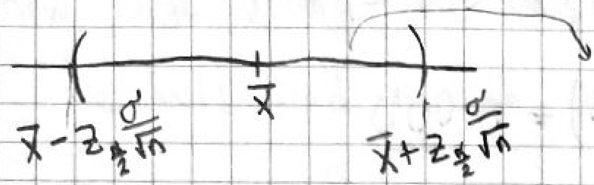
$$P\left(-Z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-Z_{\alpha/2} < \frac{\mu-\bar{X}}{\sigma/\sqrt{n}} < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(\bar{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

⊛ ↳ Confidence interval of population mean.

<u>Definition:</u> Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with known variance σ^2. $(1-\alpha)100\%$ confidence interval on population mean μ is given by $\boxed{\overline{X} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}$



μ is in this range with probability $(1-\alpha)$.

(*) Empirical studies show that for samples of size $n \geq 25$, the above boundaries are usually satisfactory even when X is not normal.

<u>Central Limit Theorem:</u> Let $X_1, X_2, \ldots X_n$ be a random sample from a distribution with mean μ and st. dev. σ. Then, for large n, $\overline{X}$ is approximately normal with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$. Moreover,

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \approx Z \quad \text{for } n \geq 25$$

$\boxed{\text{Ex}}$ The service time for a customer coming through a checkout counter in a store, is a random variable with mean 1.5 mins and variance of 1 min. Approximate the probability that 100 customers can be served in less than 2 hours of total service time.

X = service time for 1 customer

$\mu = 1.5 \quad \sigma^2 = 1 \quad \sigma = 1$

$P\left(\begin{smallmatrix}100 \text{ customers} \\ \text{service time}\end{smallmatrix} \leq 120\right) =$

$P\left(\frac{X_1 + X_2 + \dots + X_{100}}{100} \leq \frac{120}{100}\right) =$

$P(\overline{X} \leq 1.2) = ?$

$P(\overline{X} \leq 1.2) = P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{1.2 - 1.5}{1/10}\right)$

$\approx P(Z \leq -3) = 0.5 - P(0 \leq Z \leq 3) = \underline{0.0013}$

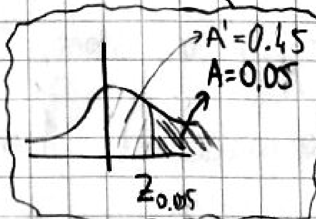
$n = 100 \geq 25 \Rightarrow$ Use CLT.

Ex Let the numbers 0, 1, 2, 9 be a random sample of four observations from a normally distributed population with a variance of 4. Construct a 90% confidence interval for the population mean μ.

$(1-\alpha)100\%$ conf. interval for μ is $\left(\overline{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$

With $\overline{X} = \frac{0+1+2+9}{4} = 3$, $\sigma = 2$, $n = 4$

$(1-\alpha)100\% = 90\% \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow Z_{\frac{\alpha}{2}} = Z_{0.05} = 1.645$



We are 90% sure that population mean is in interval $\left(3 - 1.645 \cdot \frac{2}{\sqrt{4}}, 3 + 1.645 \cdot \frac{2}{\sqrt{4}}\right)$

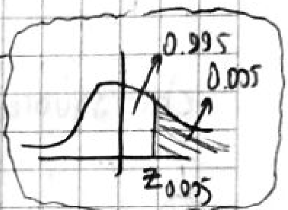
$= (1.355, 4.645)$

Ex | If we wish to be 99% sure that a random sample taken from a normally distributed population with st.dev 15, yields a mean that is within 2 units of the true mean, how large should the sample be?

We should find n such that $\mu \approx \bar{x}$, $\text{Error} = |\mu - \bar{x}| \leq 2$ with prob. 99%.

$E_{max} = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq 2$ We have $\sigma = 15$

$99\% = (1-\alpha)100\% \Rightarrow \frac{\alpha}{2} = 0.005 \Rightarrow z_{0.005} = 2.575$

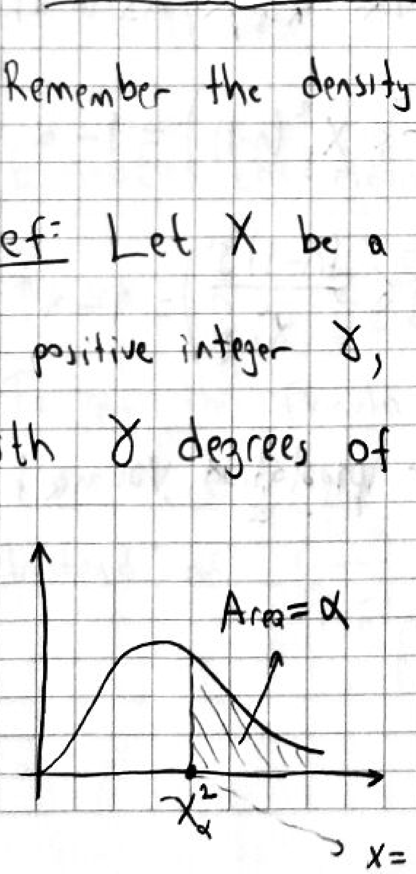


$2.575 \cdot \frac{15}{\sqrt{n}} \leq 2 \Rightarrow n \geq 372.93 \Rightarrow \underline{n \geq 373}$

Chi-Square Distribution

Remember the density function for Gamma dist is $f(x) = \frac{1}{\Gamma(\alpha)\cdot\beta^{\alpha}} \cdot x^{\alpha-1} \cdot e^{-\frac{x}{\beta}}$ for $x > 0$

Def: Let X be a Gamma rand. var with $\beta = 2$ and $\alpha = \frac{\gamma}{2}$ for a positive integer γ, is said to have a chi-squared distribution with γ degrees of freedom. We will denote this with χ^2_{γ}.



when d.f. $\gamma = 15$

$\chi^2_{0.100}(15) = 22.307$

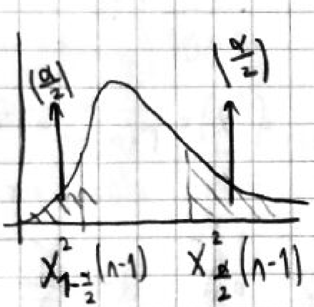
__Theorem:__ Let s^2 be the sample variance based on a random sample of size n from a distribution with mean μ, var σ^2. Then, s^2 is an unbiased estimator of σ^2.

__Proof:__ We shold show $E[s^2] = \sigma^2$.

__Theorem:__ Let $X_1, X_2, \dots, X_n$ be a random sample from a distribution with mean μ, variance σ^2. Then,

The random variable $\frac{(n-1)}{\sigma^2} s^2 = \frac{(n-1)}{\sigma^2} \sum_{k=1}^{n} \frac{(X_k - \overline{X})^2}{n-1}$ has a chi-squared distribution with $(n-1)$ degrees of freedom.

⊛ Let X^2_{n-1} be chi-squared random variable with $(n-1)$ degrees of fr.



We have $P\left(X^2_{1-\frac{\alpha}{2}}(n-1) \leq X^2_{n-1} \leq X^2_{\frac{\alpha}{2}}(n-1)\right) = 1-\alpha$

Then, $P\left(X^2_{1-\frac{\alpha}{2}}(n-1) \leq \frac{(n-1)s^2}{\sigma^2} \leq X^2_{\frac{\alpha}{2}}(n-1)\right) = 1-\alpha$

, $P\left(\frac{(n-1)s^2}{X^2_{\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-1)s^2}{X^2_{1-\frac{\alpha}{2}}}\right) = 1-\alpha$

__Def:__ $(1-\alpha)100\%$ confidence interval for population variance;

$$\left(\frac{(n-1)s^2}{X^2_{\alpha/2}}, \frac{(n-1)s^2}{X^2_{1-\alpha/2}}\right)$$

similor
HW3-Q4

not used

Ex] 12 bank tellers were randomly sampled and it was determined that they made average of 3.6 errors per day with a sample st. deviation of 0.42 errors. Assume that the number of errors bank tellers make per day is normally distributed. Construct a 90% confidence interval for the population variance of the number of errors per day.

$n = 12$, $90\% = (1-\alpha)100\% \Rightarrow \frac{\alpha}{2} = 0.05$

$X^2_{\frac{\alpha}{2}} = X^2_{0.05} = 19.675$, $X^2_{1-\frac{\alpha}{2}} = X^2_{0.95} = 4.575$

90% confidence interval for population for δ^2 is,

$$= \left(\frac{11 \cdot (0.42)^2}{19.675}, \frac{11(0.42)^2}{4.575} \right) = (0.0986, 0.4241)$$

Estimating the mean and student-t distribution

We use $\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ as confidence interval for μ when δ^2 is known.

To get this formula we used the fact $\frac{\bar{X} - \mu}{\delta/\sqrt{n}}$ when X is normal.

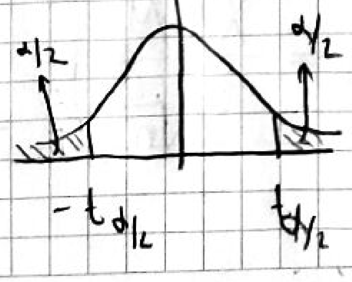
Instead of $\frac{\bar{X} - \mu}{\delta/\sqrt{n}}$, let us consider $T = \frac{\bar{X} - \mu}{s/\sqrt{n}}$

Theorem: Let $X_1, X_2, \dots, X_n$ be a random sample from a normal distribution with mean μ and variance σ^2. The random variable $\left(\frac{\bar{X}-\mu}{s/\sqrt{n}}\right)$ follows a T distribution with $(n-1)$ degrees of freedom

where $$T = \frac{Z}{\sqrt{\chi^2_{n-1}/(n-1)}}.$$

Properties of T distribution

1. There are infinetly many T dists, each defined by one parameter γ, called degrees of freedom.
2. Each T random variable is continuous random variable with density given by $$f(t) = \frac{\Gamma((\gamma+1)/2)}{\Gamma(\gamma/2)\cdot\sqrt{\pi\gamma}} \cdot \left(1+\frac{t^2}{\gamma}\right)^{-\frac{\gamma+1}{2}}, \quad -\infty < t < \infty$$
3. The graph of the density function is a symetric bell-shaped curve.
4. As the number of degrees of freedom increases, the bell shaped curve for T_γ approachs the standard normal curve.



$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1-\alpha$$

$$P\left(-t_{\alpha/2} < \frac{\bar{X}-\mu}{s/\sqrt{n}} < t_{\alpha/2}\right) = 1-\alpha$$

$$P\left(\bar{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right) = 1-\alpha$$

Def: $(1-\alpha)100\%$ confidence interval for μ is $\left(\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}\right)$, provided that sample is taken from a normal population.

Ex) A sample of 12 measurements resulted in a mean 585 and a st. dev. of 38. Estimate the true mean using 95% confidence interval. List any assumptions you need to make.

$\bar{X} = 585$, $S = 38$, $n = 12$

$95\% = 100\%(1-\alpha) \Rightarrow \frac{\alpha}{2} = 0.025$

$\Big\}$ $t_{\frac{\alpha}{2}} = t_{0.025} = 2.201$

$d.f = n-1 = 11$

95% conf. interval for μ is $\left(585 - 2.201 \cdot \frac{38}{\sqrt{12}}, 585 + 2.201 \cdot \frac{38}{\sqrt{12}}\right)$

$= (560.8558, 609.1442)$

We assumed that our sample came from a population with normal dist.

HW3-Q3) Is $\hat{\theta}_3$ unbiased? $E[\hat{\theta}_3] = aE[\hat{\theta}_1] + (1-a)E[\hat{\theta}_2]$

$\overset{\circledast}{=} a\theta + (1-a)\theta = \theta$ $\circledast$: Since θ_1, θ_2 are unbiased.

If $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent, how should a be chosen to min. $V(\hat{\theta}_3)$.

$V(\hat{\theta}_3) = V(a\hat{\theta}_1 + (1-a)\hat{\theta}_2) \overset{\boxed{*}}{=} a^2\sigma_1^2 + (1-a)^2\sigma_2^2$ $\boxed{*}$: Since $\hat{\theta}_1, \hat{\theta}_2$ ind.

$$\frac{dV(\hat{\theta}_3)}{da} = 2a\sigma_1^2 + 2(a-1)\sigma_2^2 = 2\left(a(\sigma_1^2 + \sigma_2^2) - \sigma_2^2\right) = 0 \Leftrightarrow a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

This a is a abs. min since $y = V(\hat{\theta}_3)$ is a up-looking parabola.

Note that $0 \leq a \leq 1$ is satisfied.

HW3-Q5 (a) $L(x_1, x_2, \ldots, x_n) = (\alpha+1)^n x_1^{\alpha} x_2^{\alpha} \cdots x_n^{\alpha}$

$\ln L = n \ln(\alpha+1) + \alpha \ln(x_1 \cdot x_2 \cdots x_n)$

$\frac{d \ln L}{d\alpha} = \frac{n}{\alpha+1} + \ln(x_1 x_2 \cdots x_n) = 0 \Rightarrow \hat{\alpha} = -1 - \frac{n}{\ln(x_1 x_2 \cdots x_n)}$

(b) $E[X] = \int_0^1 (\alpha+1) x^{\alpha+1} \, dx = \frac{\alpha+1}{\alpha+2} \quad M = \frac{\alpha+1}{\alpha+2} \Rightarrow \hat{M} = \frac{\hat{\alpha}+1}{\hat{\alpha}+2} = \ldots$