改善

KAIZEN

TEAM

Ch 1

Statistics: Branch of mathematics that deal with:

→ Collection: ① retrospective ② observational (→ observe & record w/out interaction) ③ design of expirement

→ presentation

→ Analysis / interpretation: To solve / make decision / design product / process

⇒ Data
- → Discrete
  - → counts
  - → nominal
  - → ordinal
- → Continous
  - → interval
  - → ratio

count, nominal, ordinal, interval, ratio ↑ have all of the above

"categorical, numerical" } nominal

if there's data before 0 → interval

if 0 = thing disappear → ratio not interval

Population: well defined set of objects

Greek letters

- pop. parameter: measure calculated to describe a characteristic for pop.
  - ↳ central tendency (μ)
  - ↳ variation (δ^2)

Latin letters

- Sample stat.: a measure ... for the sample, with/without replacment

1) Random (unbiased): Every element has equal chance of being selected

2) Biased: prefer an element

x_1, x_2, x_n — pop.

Sample — subset of pop.

Statistic → Descriptive

↳ pop. par. → unknown → take sample → calculate sample stat. → pop. par.

- Probability
  - → Pop. par. → known
  - → Sample → unknown

Ch 2 Probabilities

Random experiment: process of obtaining observations or measurments

- Ex.: can result in different outcomes eventhough it is repeated in the same manner outcomes cant be predicted with certainty

- Sample space: set of all possible outcomes of R.E
  - ↳ Discrete S.S: it consists of finite / infinite countable set of outcomes
  - ↳ Continous S.S: it contains an interval (finite / infinite) of real numbers

- Event: any collection (subset) of outcomes contained in S.S
  - ↳ simple → only 1 outcome // compound → more than 1 outcome

---

Relation from set theory

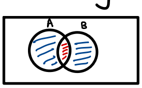
$\boxed{1}$ Union : The union of two events A & B [U : OR]
all elements in A or in B or in both [A U B]

$\boxed{2}$ intersection: event consisting of all outcomes that are in both A & B
[∩ : AND] [A ∩ B]

$\boxed{3}$ Mutually exclusive events : no outcomes in common (Disjoint)

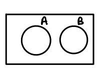
$\boxed{4}$ complement : all outcomes in the S.S but not in A [$\bar{A}$, A', not A]

Venn Diagrams



AUB

AUB



M.E

DeMorgan Law's

$\overline{A \cup B} = \overline{A} \cap \overline{B}$

$\overline{A \cap B} = \overline{A} \cup \overline{B}$

Probability : Quantity used to quantify chance that an event (outcome) of a R.E will occur

1. $0 \leq P(A) \leq 1$ * $P(\bar{A}) = 1 - P(A)$

2. $P(S) = 1$

3. if $A_1, A_2 \dots$ are sequence of mutually exclusive events then,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

P of union = P of summation

* previous probabilities → they dont tell us how to determine Prob.,

to determine prob. :-

$S = \{$ Sun, Mon, Tue, Wed, Thu, Fri, Sat $\}$

① Evaluation of possible outcomes

$P(M) = \frac{1}{7}$ $P(T) = \frac{2}{7}$ $P(D) = \frac{3}{7}$ $P(S) = \frac{2}{7}$ $P(\bar{M}) = \frac{6}{7}$

② Empirical evidence (experiment)

$E_1 = T \cup S \cup D$

$= \frac{2}{7} + \frac{2}{7} + \frac{3}{7} = \frac{7}{7} = 1$

$E_2 = M \cup T \cup S \cup D$

$= \frac{1}{7} + \frac{2}{7} + \frac{2}{7} + \frac{3}{7} = \frac{8}{7}$ ✗

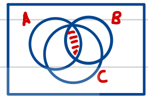
$\frac{8}{7} - \frac{1}{7} = 1 \left(\frac{7}{7}\right)$

mutual between M & D

Addition rule

mutual

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



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

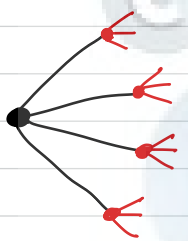
Counting techniques

Product

① Multiplication rule: if an operation can be described as a sequences of k steps, & the number of ways of counting: step 1 in n_1, 2 in n_2, ... k in n_k then the total number of ways of completing the operation is $n_1 * n_2 * \dots n_k$

Ex A company wants to ship: 1) select airline } steps
2) select airplane

steps → $k = 2$, $n_1 = 4$ airlines , $n_2 = 3$ types



$n_1 * n_2 =$

$4 * 3 = 12$

② Combinations: given a set of n distinct (different) objects any unordered subset of size r objects is called combinations

(Choosen r from n)

$$^{n}C_r = \frac{n!}{(n-r)!\,r!}$$

* F, M, L ⟹ n=3 The order is not imp in com.

C = FM / ML / FL

③ Permutations: Any order sequences of r object from a set of n distinct object is called permutation, The order matters

$$^{n}P_r = \frac{n!}{(n-r)!}$$

* F, M, L ⟹ n=3 The order is imp. in per.

P = FM / MF / ML / LM / FL / FL

$\boxed{P > C}$

④ Permutation of similar objects: number of permutations of n objects in which $n = n_1 + n_2 + \dots + n_r$ of which $n_1, \dots, n_r$ are $1^{st}, \dots, r^{th}$ type

* 5 students $n_1 = 3$ Females, $n_2 = 2$ males

$$P_{similar} = \frac{n!}{n_1!\,n_2! \cdots n_r!}$$

$$P = \frac{5!}{3!\,2!} = \frac{5*4*\cancel{3!}}{\cancel{3!}\,2*1} = \frac{20}{2} = 10$$

, 10 ways to order the 5 students

FFFMM / FMMFF / FMFMF / FFMMF / MFMFF / MMFFF / FMFFM / MFFFM / FFMFM / MFFMF

A package of 7 light bulbs contains 2 defectives if 3 bulbs are select at rand.

$\boxed{1}$ whats the total # of ways we can select the bulbs?

$$C^7_3 = \frac{7!}{(4)!\,3!} = \frac{7 \cancel{*} \cancel{6} \cancel{*} 5 \cancel{*} \cancel{4!}}{\cancel{4!} \cancel{*} \cancel{6}} = 35$$

$\boxed{2}$ How many ways we can select one defective?

2 ways

$\boxed{3}$ what is the prob. that we choose 1 defective bulb?

$$\frac{C^2_1 \times C^5_2}{C^7_3} = \frac{2 \times 10}{35} = \frac{20}{35}$$

$$C^5_2 = \frac{5!}{3!\,2!} = \frac{5 \times 4 \times \cancel{3!}}{\cancel{3!}\;2} = 10$$

if 2:

$$\frac{C^2_2\;C^5_1}{C^7_3} = \frac{1 * 5}{35} = \frac{1}{7}$$

A package of 10 bulbs 4 are defectives

1. What is probability we find 1 defective?
2. " " " " " 2 defectives?

assuming → 1 D | 3 ND

10 → 4 D, 6 ND

$$P(1D) = \frac{C^4_1 \cdot C^6_3}{C^{10}_4}$$

$$P(2D) = \frac{C^4_2 \cdot C^6_2}{C^{10}_4}$$

② Empirical evidence: frequency relative

$0 \leq P(A) \leq 1$

contingency table

	H	L	
M	28	35	63
$\bar{M} \equiv F$	42	31	73
	70	66	136

equivalent

① what is the pro. that the selected student scored high?

$P(H) = 70/136$

② what is the pro. that the selected student is male?

$P(M) = 63/136$

③ what is the pro. the the selected student is male & scored high?

$P(M \cap H) = 28/136$

④ what is the prob. that the selected is male or high?

$$P(M \cup H) = P(M) + P(H) - P(M \cap H) = \frac{60 + 70 - 28}{136}$$

⑤ what is the pro. ... is female or scored low?

$$P(F \cup L) = P(\bar{M} \cup \bar{H}) = P(\overline{M \cap H}) = 1 - P(M \cap H) = \frac{136 - 28}{136}$$

$\boxed{\text{2-5 conditional pro.}}$

⑥ What is the pro. that the selected male student scored high?

$$P(H \mid \underset{\text{given}}{M}) = 28/63 = \frac{P(H \cap M)}{P(M)} = \frac{28/\cancel{136}}{63/\cancel{136}}$$

2-5 conditional pro.

Examine how the information "an event B occurred" affects the pro. assigned to event A $\rightarrow P(A|B) = \frac{P(A \cap B)}{P(B)}$, $P(B) > 0$

	H	L	
M	32	31	63
F	38	35	73
	70	66	136

① what is the pro.... scored high?

$P(H) = 70 / 136 = 0.51$

② what is the pro. ... is male?

$P(M) = 63 / 136$

③ what is the pro. ... is male & scored high?

$P(M \cap H) = 32 / 136$

last ex.
⑥
④ what is the pro. ... that the selected male student scored high?

$P(H|M) = 32/63 = 0.51$

2-7 independent events

* Knowledge that an outcome of an exp. is in event B does not affect the pro. that an outcome is in A } ind.

* Events A & B are independent events, so :-

A given B   given

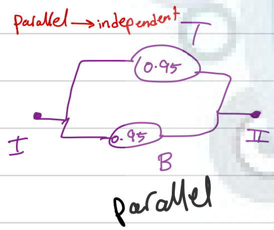
1) $P(A|B) = P(A)$

B given A

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

3) $P(A \cap B) = P(A) \cdot P(B)$

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



Pro. the current moves

$P(T \cup B) = P(T) + P(B) - P(T \cap B) = 0.95 + 0.95 - (0.95)^2$

$P(T \text{ or } B)$

$P(\bar{T} \cap \bar{B}) = 1 - P(T \cup B)$

* $P(T \cup B) = 1 - P(\overline{T \cup B}) = 1 - P(\bar{T} \cap \bar{B})$

$P(\bar{T} \cap \bar{B}) = P(\bar{T})\, P(\bar{B})$

$P(T \cup B) = 1 - P(\bar{T} \cap \bar{B}) = 1 - (P(\bar{T})\, P(\bar{B}))$

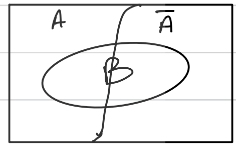
series

$P(L \& R) = P(L \cap R) = P(L)\, P(R)$

2-6 Multiplication & total Pro. Rule

$P(A|B) = \dfrac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$

$B = A \cap B \cup \bar{A} \cap B$

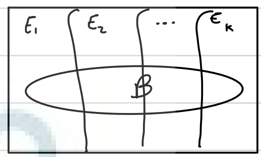


$P(B) = P((A \cap B) \cup (\bar{A} \cap B))$

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

$P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})$

$E_1, E_2, \dots, E_k \rightarrow$ Mutually exclusive



$E_1 \cup E_2 \cup \dots \cup E_k = S \longrightarrow$ Exhaustive their union gives us the sample space

$P(B) = P(E_1 \cap B) + P(E_2 \cap B) + \dots + P(E_k \cap B)$

$P(B) = P(B|E_1) \cdot P(E_1) + P(B|E_2) \cdot P(E_2) + \dots + P(B|E_k) \cdot P(E_k)$

EXAMPLE 2.27 | Medical Diagnostic

Because a new medical procedure has been shown to be effective in the early detection of an illness, a medical screening of the population is proposed. The probability that the test correctly identifies someone with the illness as positive (known as the sensitivity) is 0.99, $\rightarrow P(B|A)$ and the probability that the test correctly identifies someone without the illness as negative (known as the specificity) is 0.95. $\rightarrow P(\bar{B}|\bar{A})$ The incidence of the illness in the general population is 0.0001. $\rightarrow P(A)$ You take the test, and the result is positive. What is the probability that you have the illness?

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

$P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})$

$0.99\,(0.0001) + \underset{1-0.95}{0.05}\,\underset{1-0.0001}{(0.9999)}$

A : Person is ill

B : test result is +ve

2.8.1 WP SS VS Customers are used to evaluate preliminary product designs. In the past, 95% of highly successful products received good reviews, 60% of moderately successful products received good reviews, and 10% of poor products received good reviews. In addition, 40% of products have been highly successful, 35% have been moderately successful, and 25% have been poor products.

$P(G|H)$ $P(G|M)$ $P(G|P)$

$P(H)$ $P(M)$ $P(P)$

a. What is the probability that a product attains a good review? $P(G)$

b. If a new design attains a good review, what is the probability that it will be a highly successful product? $P(H|G)$

c. If a product does not attain a good review, what is the probability that it will be a highly successful product? $P(H|\bar{G})$ / $P(H|B)$

Product type → H, M, P

Review → G, B

$$P(G) = P(G|H)\,P(H) + P(G|M)\,P(M) + P(G|P)\,P(P)$$

$$0.95\,(0.4) + 0.6\,(0.35) + 0.1\,(0.25)$$

$$P(H|G) = \frac{P(G|H)\,P(H)}{P(G)}$$

$P(H|\bar{G})$

$P(H|B)$

2.8.4 **WP** Software to detect fraud in consumer phone cards tracks the number of metropolitan areas where calls originate each day. It is found that 1% of the legitimate users originate calls from two or more metropolitan areas in a single day. However, 30% of fraudulent users originate calls from two or more metropolitan areas in a single day. The proportion of fraudulent users is 0.01%. If the same user originates calls from two or more metropolitan areas in a single day, what is the probability that the user is fraudulent?

0.01 $P(A|\bar{F})$

0.01 $P(A|F)$

0.0001 $P(F)$

$\searrow L$

$\nearrow F$

A:

$$P(F|A) = \frac{P(A|F)\,P(F)}{P(A)}$$

2-9 Random Veriables

Toss a coin twice

$S = \{\underset{2}{HH}, \underset{1}{HT}, \underset{1}{TH}, \underset{0}{TT}\}$

X : Represents # of heads

R.V is a function that assigns a real # to each outcome in the S.S of a R.E

discrete

D.R.V : a random variable that can take on only a countable (finite / infinite) of Possible values

continous

C.R.V : a random variable with an interval (finite / infinite) of Real #s for the range

Range of R.V X : The set of possible # of a R.V X

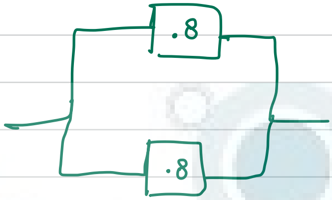
$P(\overset{X}{2}\overset{x}{H}) = 0.25$ $P(1H) = 0.5$ $P(0H) = 0.25$

Probability distribution $[P(X) = x]$ X : RV , x : general form for the value that the R.V could assume

3-1

formula
Table
Figure

x	$P(X=x)$
0	0.04
1	0.32
2	0.64



$P(\bar{T} \cap \bar{B}) \leftarrow P(X=0) = 0.04$

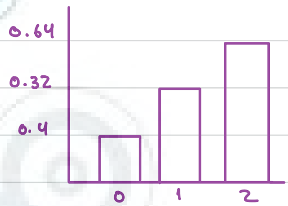
$P(X=1) = 0.32 \rightarrow P(T \cap \bar{B}) \overset{\text{or}}{\cup} P(\bar{T} \cap B)$

$(0.8)(0.2) + (0.2)(0.8)$

$0.16 + 0.16$

$P(T \cap B) \leftarrow P(X=2) = 0.64$

X : # of switches that close



Prob. distribution formula : is convenient to express the prob. in terms of formula $P(x)$ or $f(x)$ called prob. mass fun. (p.m.f)

* mass → discrete

- p.m.f for D.R.V is denoted $f(x_i)$ & represents the prob. that X take on the value x_i when x_i is possible values & $i = 1, 2, \cdots, n$ or ∞

- The p.m.f has the following properties (all of them) :-
  1) $f(x_i) = P(X = x_i)$
  2) $f(x_i) \geq 0$
  3) $\sum_{i=1}^{n \text{ or } \infty} f(x_i) = 1$ ← depends on i

Same previous example

$$f(x) = \begin{cases} 0 & x < 0 \\ 0.04 & 0 \leq x < 1 \\ 0.32 & 1 \leq x < 2 \\ 0.64 & x = 2 \end{cases}$$

* Example: determine if the functions are appropriate p.m.f ?

a) $f(x) = \frac{x-2}{2}$, $x = 1, 2, 3, 4$

$f(1) = \frac{1-2}{2} = \frac{-1}{2} < 0$, $f(x)$ is not appropriate p.m.f

b) $h(x) = x^2/25$, $x = 0, 1, 2, 3, 4$

$h(0) = 0$ $\quad$ $h(3) = 9/25$

$h(1) = 1/25$ $\quad$ $h(4) = 16/25$

$h(2) = 4/25$ $\quad \rightarrow h(x)$ is not ... $\Rightarrow \sum_{i=0}^{n=4} h(x) \neq 1$

3-2 Cumulative Distribution functions

A cumulative distribution fn (c.d.f) of d.r.v X with p.m.f $f(x)$ is defined for every number x as $F(x) = P(X \leq x) = \sum_{X \leq x} f(x)$

- $F(x)$ should satisfy the following properties:-

$0 \leq F(x) \leq 1$

if $x \leq y$, then $F(x) \leq F(y)$

same example)

x	$P(X=2)$	$F(x)$
0	0.04	0.04
1	0.32	0.36
2	0.64	1.0

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.04 & 0 \leq x < 1 \\ 0.36 & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

$F(0) = P(X \leq 0) = f(0) = 0.04$

$F(3) = P(X \leq 3) = f(2) = 1$

$F(1.5) = P(X \leq 1.5) = f(0) + f(1)$
$= 0.04 + 0.32 = 0.36$

Example:

x	$f(x)$
-2	$1/8$
-1	$2/8$
0	$2/8$
1	$2/8$
2	$1/8$

① $P(x \leq 2) = 1$

② $P(x > -2) = 7/8$

③ $P(-1 \leq x \leq 1) = 6/8$

④ $P(\underset{A}{x \leq -1} \cup \underset{B}{x = 2}) = P(A) + P(B) - \underset{M.E}{\emptyset}$

$= P(x \leq -1) + P(x = 2) - \emptyset$

$= \frac{3}{8} + \frac{1}{8} = \frac{1}{2}$

$$F(x) = \begin{cases} 0 & x < -2 \\ 1/8 & -2 \leq x < -1 \\ 3/8 & -1 \leq x < 0 \\ 5/8 & 0 \leq x < 1 \\ 7/8 & 1 \leq x < 2 \\ 1/8 & 2 \leq x \end{cases}$$

$F(3) = P(x \leq 3) = 1$

$P(x \leq 1.25) = F(1.25) = 7/8$

$P(x \leq 2.2) = F(2.2) = 1$

$P(-1.1 < x \leq 1) = F(1) - \frac{1}{8}$

$= \frac{7}{8} - \frac{1}{8} = \frac{6}{8}$

$F(-2)$

3-3 Mean of D.R.V

X

Central tendency → mean

$\mu = E(X) \rightarrow$ expected value of x

$= \sum_{i=1}^{n} x_i f(x_i)$

164	170	160	160	168
x_1	x_2	x_3	x_4	x_5

$\mu = \underbrace{\frac{x_1 + x_2 + x_3 + x_4 + x_5}{5}} = \frac{x_1 + x_2 \cdots}{n}$

$= \frac{x_1 \, 1}{n} + \frac{x_2 \, 1}{n} + \dots + \frac{x_n \, 1}{n}$

value x * prob $(\frac{1}{n})$ [weight]

Example 1 :

num of head x	$f(x)$
0	0.25
1	0.5
2	0.25

$E(x) = 0(0.25) + 1(0.5) + 2(0.25) = 1$

Example 2 :

R	x	$f(x)$
15	1.5	0.25
16	1.6	0.6
20	2	0.15

$E(x) = (1.5)(0.25) + (1.6)(0.6) + (2)(0.15) = 1.635$

$1 \rightarrow 10$ JD $\Rightarrow$ R

$E(10x) = E(g(x)) = \sum_{i=1}^{n} g(x_i) \cdot f(x_i)$

as x

$g(x) = 10x^2 + 20x - 5$

$E(g(x)) = (10(1.5)^2 + 20(1.5) - 5)\, 0.25 + (10(1.6)^2 + 20(1.6) - 5)\, 0.6 + (10(2)^2 + 20(2) - 5)\, 0.15$

$54.685 = 11.875 + 31.56 + 11.25$

* Variance of D.R.V

$E(x-E(x))^2$

$\sigma^2 = V(x) = E(X-\mu)^2$

prob

$= \sum (x-E(x))^2 \cdot f(x) = \sum x^2 f(x) - \mu^2$

shift one unit

$* \; E(x+b) = E(x) + E(b)$

$= E(x) + b$

constant

$* \; E(ax) = E(a) \cdot E(x)$

$= a \cdot E(x)$

changes

$* \; E(ax+b) = aE(x) + b$ → shift

$* \; V(ax+b) = a^2 V(x)$

$* \; V(x+b) = V(x)$ doesn't change $V(b)=0$

$* \; V(ax) = a^2 V(x)$

Example 1:

# of heads

x	$f(x)$
0	0.25
1	0.5
2	0.25

$V(X) = 0^2(0.25) + 1^2(0.5) + 2^2(0.25) - 1^2 = 0.5$

3-4 Discrete Uniform Distribution

$f(x) = \frac{1}{n}$ $\qquad\qquad$ $n = b - a + 1$

A R.V X has a discrete uniform distribution if each of the n values in its range say $x_1, x_2, \ldots, x_n$ has equal prob.

Consecutive integers:

* $a, a+1, a+2, \ldots, b$

$$\mu = E(x) = \frac{b+a}{2}$$

$$\sigma^2 = V(x) = \frac{(b-a+1)^2 - 1}{12} = \frac{n^2-1}{12}$$

* $a, a+k, a+2k, a+3k, \ldots, b$

$$E(x) = \frac{b+a}{2}$$

$$V(x) = \frac{k^2(n^2-1)}{12}$$

Bernoulli Distribution

3-5 Binomial Distribution

X : any R.V whose only possible values are 0 or 1 (Bernoulli)

$f(x) = p^x (1-p)^{1-x}$ for $x = 0, 1$ $\quad f(x) = P(X=x)$

$\mu = E(x) = p$

$\sigma^2 = V(x) = p(1-p)$

x	$f(x)$
0	$1-p$
1	p

$f(1) = P(X=1) = p^1 (1-p)^{1-1} = p^1 (1-p)^0 = p$

Binomial

n: fixed # of Bernoulli R.V $\quad f(x) = C^n_x \, p^x (1-p)^{n-x}$ for $x = 0, 1, \dots, n$

$f(x) = p^x$

p: prob. of event is constant

success (my target)

$P(M) = 0.4$ $\quad M = \frac{20}{50}$ $\quad$ X : # of males $\rightarrow 0, \dots, n$

$P(F) = 0.6$

$n = 5$ $\quad$ success $\rightarrow M$

$P(X=0) = P(F \cap F \cap F \cap F \cap F)$ $\quad$ events are independent

$= P(F) \cdot P(F) \cdot P(F) \cdot P(F) \cdot P(F) = (1-p)^{n-x}$

$P(X=5) = p^5 = p^x$

Binomial Dist.

X: total # of success in n trials

1. Bernoulli trials
2. independent
3. p is const.
4. fixed $\underline{\underline{n}}$

$$f(x) = C^n_x \, p^x (1-p)^{n-x} \quad , \ x = 0, 1, \dots, n$$

$$E(x) = n \cdot p$$

$$V(x) = n \cdot p (1-p)$$

Binomial → success
→ independent

- The phone lines to an airline reservation system are occupied 40 % of the time. Assume that the events that the lines are occupied on successive calls are independent.
- Assume that 10 calls are placed to the airline.

1. What is the probability that for exactly three calls the lines are occupied?
2. What is the probability that for exactly one call the lines are occupied?
3. What is the probability that for exactly one call the lines are not occupied? $\rightarrow \bar{p} = 1 - p = q$ > same
4. What is the probability that for exactly nine calls the lines are occupied? $p = q = \bar{p} = 1$
5. What is the probability that for at least one call the lines are not occupied?
6. What is the expected number of calls in which the lines are all occupied?

$P(B) = 0.4$

$P(\bar{B}) = 0.6$

$n = 10$

1) 0.215

2) 0.040

3) 1.57286

4) 1.57286

6) $E(B) = n \cdot p = 4$

5) $P(\bar{B} \geq 1) = 1 - P(\bar{B} < 1) = 1 - P(\bar{B} = 0)$

imp.

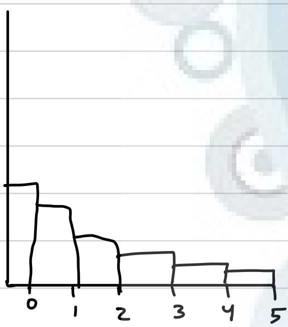
or

$P(B \leq 9) = 1 - P(B > 9) = 1 - P(B = 10) = 0.9999$

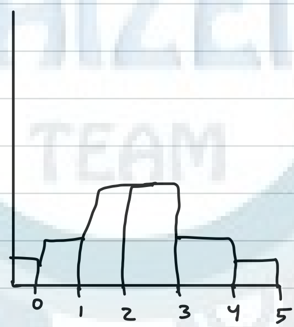
$n = 5$	$n = 5$	$n = 5$
$p = 0.1$	$p = 0.5$	$p = 0.9$
0.59	0.03125	0.00001
0.328	0.15625	0.00045
0.0729	0.3125	0.0081
8.1×10^{-3}	0.3125	0.0729
4.5×10^{-4}	0.15625	0.32805
1×10^{-5}	0.03125	0.59049

$f(0)$	$\cancel{{}^5C_0}\,\cancel{(0.5)^0}\,(1-0.5)^5 = 0.031$
$f(1)$	$5(0.5)(0.5)^4 = 0.156$
$f(2)$	$10(0.5)^2(0.5)^3 = 0.312$
$f(3)$	0.312
$f(4)$	0.25
$f(5)$	0.031

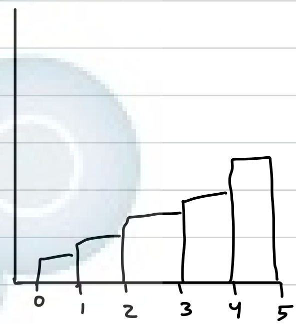
Me



skewed to the right



symmetrical



skewed to the left

3-6 Geometric & Negative Binomial Distributions

Geometric Distributions

X: # of trials until first success

$f(x) = (1-p)^{x-1} \cdot p$, for $x = 1, \ldots$

$E(x) = \frac{1}{p}$

$V(x) = \frac{(1-p)}{p^2}$

Negative Binomial Distribution

X : # of trials until the r^{th} success

$f(x) = C^{x-1}_{r-1} (1-p)^{x-r} \cdot p$, for $x = r, r+1, \ldots$

$E(x) = r \cdot \frac{1}{p}$

$V(x) = \frac{r(1-p)}{p^2}$

Binomial → defective
Poisson → defects

3-8 Poisson Distribution

X : # of events in an interval

$\underbrace{\text{events}}_{\lambda}$ $\underbrace{\text{interval}}_{t}$

$\mu = E(x) = \lambda$

$$f(x) = P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \quad , x = 0, 1, \ldots$$

$E(x) = \lambda t$ 0.59

~ same

$V(x) = \lambda t$ 0.6

① prob of no accident in a week

$$P(X=0) = \frac{(3 \times 1)^0 e^{-(3 \times 1)}}{0!} = 0.04979$$

② prob of 2 accidents in a week $t=1$

③ Prob of 2 〃 〃 two weeks $t=2$

④ 〃 〃 〃 〃 〃 half a week $t=\frac{1}{2}$

3-7 Hypergeometric Distribution

Pop. $N \begin{cases} S : K \\ F : N-K \end{cases}$ $\qquad$ S-S $n \begin{cases} x \\ n-x \end{cases}$

$$f(x) = \frac{C^{K}_{x} \cdot C^{N-K}_{n-x}}{C^{N}_{n}}$$

$$E(x) = n \cdot p \quad , \quad p = \frac{K}{N}$$

$$V(x) = n \cdot p(1-p)\left[\frac{N-n}{N-1}\right]$$

finite population correction factor

An experiment is conducted to select a catalyst used in production of soap, suppose an IE selects 3 (n) catalyst for testing from among a group of 10 (N) catalyst, six (K) of which have low acidity

$N \begin{cases} \text{success} \to K \\ \text{Failure} \to N-K \end{cases}$

$n \begin{cases} x \\ n-x \end{cases}$

① Find the prob. that one (x) low acidity catalyst is selected

G: low $\quad \frac{6}{10}$

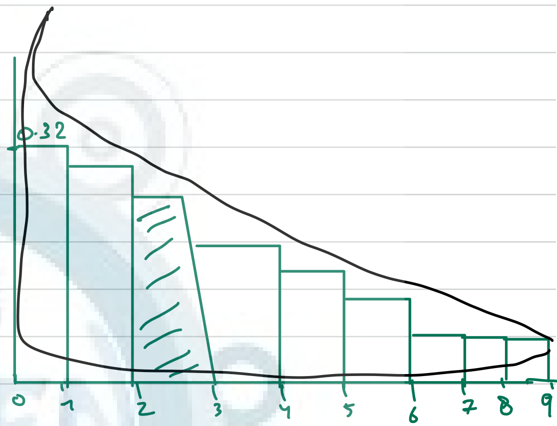
$$\frac{C^{6}_{1} \times C^{4}_{2}}{C^{10}_{3}}$$

Ch 4

Continuous R.V Distribution

50 batteries :

life	frequency	relative frequency
0 - 1	16	0.32
1 - 2	11	0.22
2 - 3	9	0.18
3 - 4	6	0.12
4 - 5	3	0.06
5 - 6	2	0.04
6 - 7	1	0.02
7 - 8	1	0.02
8 - 9	1	0.02



Area of the bin = prob.

Area under the curve = probability

p.m.f D.R.V	p.d.f C.R.V
$E(x) = \sum f(x)$	$E(x) = \int f(x)$
$V(x) = \sum x^2 f(x) - \mu^2$	$V(x) = \int x^2 f(x) - \mu^2$

$f(x)$: probability density function (p.d.f)

$\boxed{4-1}$

f(x) is called prob. density fn. for C.R.V a p.d.f in a function such that :

1) $f(x) \geq 0$ for all x — prob. can't be -ve

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

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

4) $p(X=a) = p(a \leq X \leq a) = \int_a^a f(x)\, dx = 0$ — no equal prob. in continous

$\rightarrow p(X \leq a) = p(X < a)$

The cumulative distribution of C.R.V X : $\boxed{4-2}$

$$F(x) = p(X \leq x) = p(X < x) = \int_{-\infty}^{x} f(x)\, dx$$

$$f(x) = \frac{d\,F(x)}{dx}$$

Expected value : $\boxed{4-3}$

$$\mu = E(x) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Expected value of $g(x)$:

$$E(g(x)) = \int_{-\infty}^{\infty} g(x) \cdot f(x)\, dx$$

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

$$f(x) = \begin{cases} 3x^2 & 0 \leq x \leq b \\ 0 & \text{elsewhere} \end{cases}$$

① find b so that $f(x)$ is a proper p.d.f

$$\int_0^b 3x^2 = 1$$

$$\int_{-\infty}^{0} \text{zero} + \int_0^1 + \int_1^{\infty} \text{zero}$$

$b = 1$

② find μ & σ^2 of x

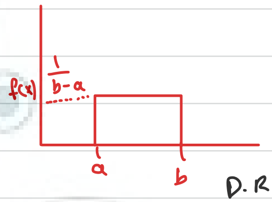
$$\mu = \frac{3}{4}$$

$$\sigma^2 = 0.0375$$

4-4 Continuous Uniform Distribution

a C.R.V has a uniform distribution of it's values spread evenly over the range of probabilities

→ Graph ⟹ Rectangular shape



D.R.V

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

$f(x) = \frac{1}{n} = \frac{1}{b-a+1}$

$\mu = E(x) = \frac{a+b}{2}$ (a, b : interval)

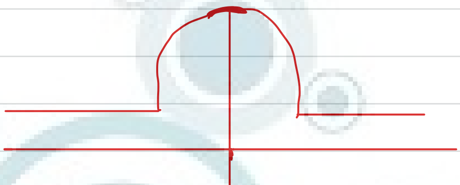
$E(x) = \frac{a+b}{2}$

$\sigma^2 = V(x) = \frac{(b-a)^2}{12}$

$V(x) = \frac{n^2-1}{12} = \frac{(b-a+1)^2-1}{12}$

4-5 Normal Distribution

- symmetrical about the mean
- Bell shape



Mean (highest probability)

Prob: Area under the curve (p.d.f)

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

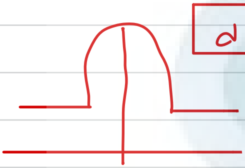
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad , \; -\infty < x < \infty$$

z

- standard normal distribution: $\mu = 0$, $\sigma = 1$ $\rightarrow$ with z not x

Transformation function: $\frac{(x-\mu)}{\sigma}$ ($z = $) to convert it to standard normal dist. from x to z

$\boxed{\sigma = 1}$

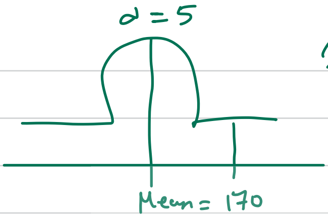


Mean

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

(Standard normal R.V)

, z: How many standard deviation are you away from the mean?



$x = 155$

$$z = \frac{155 - 170}{5}$$

$$p(x > 155) = p(z > -3) = p(z < 3) = 0.998650$$

$$p(x < 175) = p(z < 1) = 0.841345$$

$$p(x < 180) = p(z < 2) = 0.977250$$

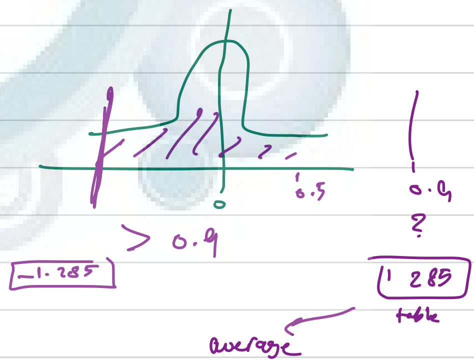
$$p(165 < x < 175) = p(-1 < z < 1) = 0.68269$$

$$p(x > 172.5) = p(z > 0.5) = 0.308538$$

$$p(x < 168) = p(z < -0.4) = 0.344578$$

$$X \longrightarrow z \xrightarrow{\text{Table}} \text{prob.}$$

$$p(z > ?) = 0.9$$



4.5.10 In 2002, the average height of a woman aged 20–74 years was 64 inches [with an increase of approximately 1 inch from 1960] (*http://usgovinfo.about.com/od/healthcare*). Suppose the height of a woman is normally distributed with a standard deviation of 2 inches.

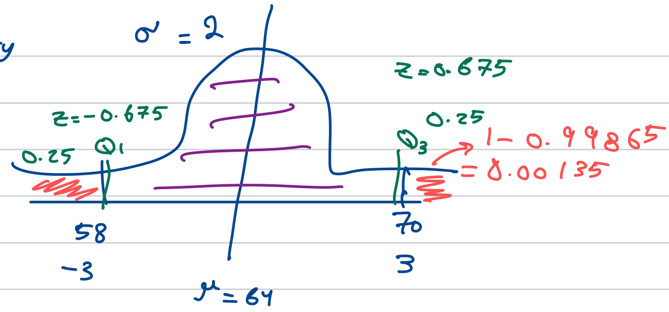
extra unnecessary info

a. What is the probability that a randomly selected woman in this population is between 58 inches and 70 inches?

b. What are the quartiles of this distribution?

c. Determine the height that is symmetric about the mean that includes 90% of this population.

d. What is the probability that five women selected at random from this population all exceed 68 inches?



a. $P(58 < x < 70) = P(-3 < z < 3) = 0.99865 - 0.00135 = 0.9975$

transformation function

$P(z < 3) = 0.99865$

b. Q_1: 1st Quartile , Q_2: 2nd quartile , Q_3: 3rd Quartile

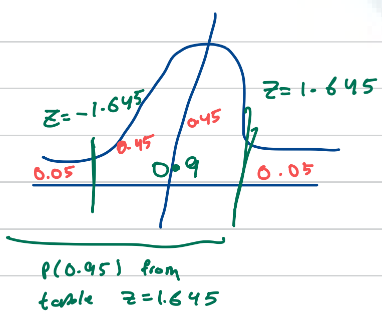
$$\frac{x-\mu}{\sigma} = z = \frac{Q_3 - \mu}{\sigma}$$

$$0.675 = \frac{Q_3 - 64}{2}$$

$\rightarrow Q_3 = 65.35 > 64$ ✓

$Q_1 = 62.65 < 64$ ✓

c.



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

$1.645 = \dfrac{x-\mu}{\sigma} \rightarrow x = 67.29$

$-1.645 = \dfrac{x-\mu}{\sigma} \rightarrow x = 60.71$

d. $z = \dfrac{x-\mu}{\sigma}$

$$\frac{68-64}{2} = 2$$

$P(z > 2) = 1 - P(z < 2)$

$1 - 0.97725$

$= 0.02275$

$P^5 = (0.02275)^5 =$

$= 6.094 \times 10^{-9}$

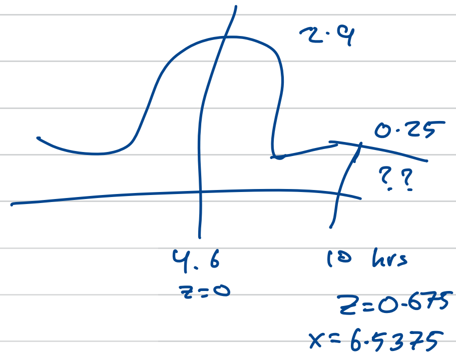
4.5.16 The length of stay at a specific emergency department in Phoenix, Arizona, in 2009 had a mean of 4.6 hours with a standard deviation of 2.9. Assume that the length of stay is normally distributed.

a. What is the probability of a length of stay greater than 10 hours?

b. What length of stay is exceeded by 25% of the visits?

c. From the normally distributed model, what is the probability of a length of stay less than 0 hours? Comment on the normally distributed assumption in this example.

$\mu = 4.6 \qquad \sigma = 2.9$



a. $p(x > 10) \rightarrow z = \frac{x-\mu}{\sigma} = \frac{10-4.6}{2.9} = 1.86$

$p(z > 1.86) = 0.968557 \rightarrow p(z > 1.86) = 1 - 0.968557 = 0.031443$

c. $z = -3 \rightarrow x = ??$