

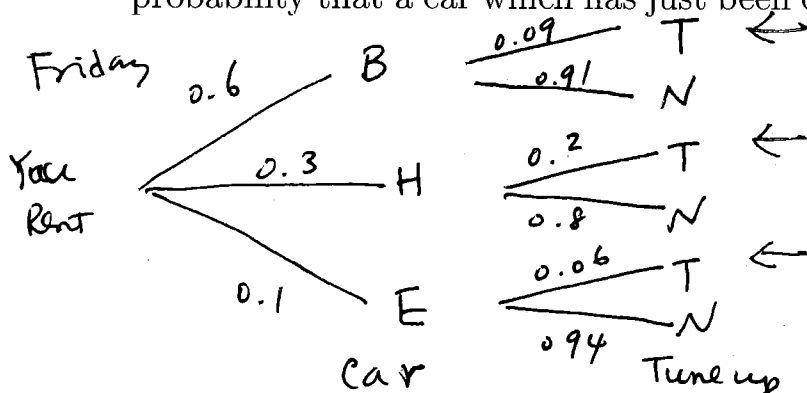
$$\begin{aligned} P(A_j \cap B) &= P(A_j | B) \cdot P(B) \\ &= P(B | A_j) \cdot P(A_j) \end{aligned}$$

THEOREM (BAYES' THEOREM) Let A_1, A_2, \dots, A_k be a collection of k mutually exclusive and exhaustive events with prior probabilities

$P(A_i)$ ($i = 1, \dots, k$). Then for any other event B for which $P(B) > 0$, the posterior prob of A_j given B is $P(A_j | B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B | A_j) \cdot P(A_j)}{\sum_{i=1}^k P(B | A_i) \cdot P(A_i)}$

Law of Total Prob

EXAMPLE Engineering Capilano rents cars from 3 rental agencies: 60% from Budget, 30% from Hertz, 10% from Eurocar. If 9% of Budget cars need a tune-up, 20% from Hertz, 6% from Eurocar, what is the probability that a car which has just been delivered will need a tune-up?



$$P(T) =$$



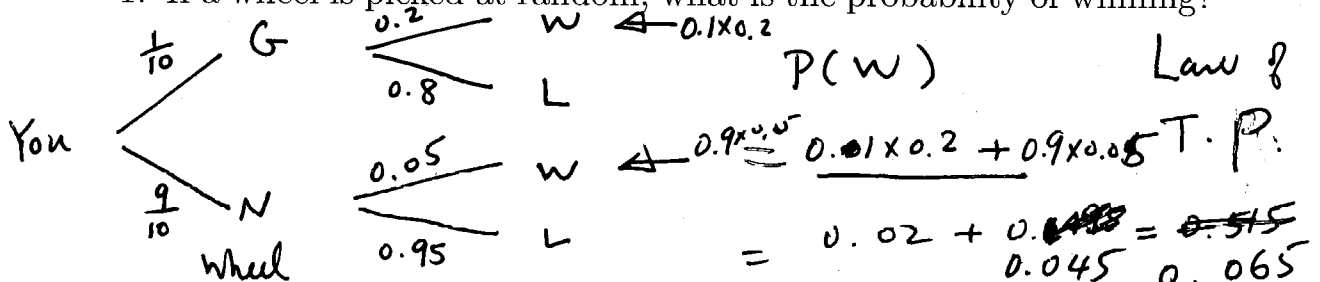
If a new rental car needs a tune-up, what is the probability that the car came from Hertz?

$$P(H | T) = \frac{P(H \cap T)}{P(T)} = \frac{P(T | H) \cdot P(H)}{0.2 \times 0.3}$$

David Hume : Problem of Induction

EXAMPLE Casino Vegas has 10 roulette wheels, indistinguishable to the customer, but one wheel (the *good wheel*) is *out of adjustment* and gives a better chance of winning (20%) than the remaining 9 wheels (5% chance).

1. If a wheel is picked at random, what is the probability of winning?



2. Having won, what is the probability that you are playing on the *good wheel*?

Fr

$$P(G|W) = \frac{P(G \cap W)}{P(W)}$$

$$= \frac{0.02}{0.065} \approx \frac{20}{65} = \frac{4}{13} \approx 0.307$$

3. Having won the first time, what is the probability of winning if you play again on the same wheel?

↓
To be continued.

$\frac{0.02}{0.065} P(G|W) \xrightarrow{0.2} W$
 $\frac{0.045}{0.065} P(N|W) \xrightarrow{0.05} W$

$$P(W|W_{\text{same wheel}}) = \frac{0.2 \times 0.02}{0.065} + \frac{0.05 \times 0.045}{0.065}$$

$$= \frac{0.004 + 0.00225}{0.065} = \frac{0.00625}{0.065} \approx 0.0965$$

4. If you keep winning, and you stay on the same wheel, what value do your *updated* probabilities of winning the next time approach?

$\xrightarrow{W \uparrow \text{ on same wheel }} P(G) = 1$
 46

$$P(W) = 0.2$$

EXAMPLE Suppose buildings collapse due to one of several errors, C_1 , C_2 , ..., C_m . For example

C_1 Poor design (underestimate load; underestimate wind stress; ...)

C_2 Poor construction (low grade materials; insufficient controls; gross error; ...)

C_3 A combination of C_1 and C_2

C_4 Other (non-assignable) causes.

From either past studies (actual percentages of buildings that have the above problems) or the subjective beliefs of expert consultants, one may determine $P(C_i)$, the *prior probabilities*.

Also, suppose that previous experience (or, again, perhaps experts' opinion) determine what the probabilities of collapse would be given each of the above cases; that is, the values of $P(B | C_i)$ are known, where B is "building collapse." These conditional probabilities are sometimes called *risk factors*.

Table 2.3 on the next page summarizes these known probabilities.

Cause C_i	Prior probability $P(C_i)$	Risk factor $P(B C_i)$	Posterior probability $P(C_i B)$
C_1	0.00050	0.10	
C_2	0.00010	0.20	
C_3	0.00001	0.40	
C_4	0.99939	0.0001	

Table 2.3: Causes of building collapse

1. Calculate the posterior probabilities.
2. What is the chance of finding an assignable cause for the collapse?
3. What is the most likely assignable cause for the collapse?
4. How likely is it that the collapse was caused by more than one assignable cause?

Practice exercises in Section 2.4: 45, 49, 53, 59, 61, 63, 67.

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

2.5 Independence

Let us first consider an example from drawing a card from a standard 52-card deck.

EXAMPLE Draw one card from a standard 52-card deck.

1. $P(K) = \frac{1}{13}$
2. $P(\spadesuit) = \frac{1}{4}$
3. $P(K|\spadesuit) = \frac{1}{13}$
4. $P(\spadesuit|K) = \frac{1}{4}$

superfluous condition.

Note that $P(K \cap \spadesuit) = \frac{1}{52} = \frac{1}{13} \cdot \frac{1}{4}$ Key idea.
However, what if we compare $P(\spadesuit | \text{Black})$ and $P(\text{Black} | \spadesuit)$?

$$P(\spadesuit | \text{Black}) = \frac{1}{2} \neq \frac{1}{4} \times \frac{1}{2} \quad P(\spadesuit) = \frac{1}{4}$$

spade

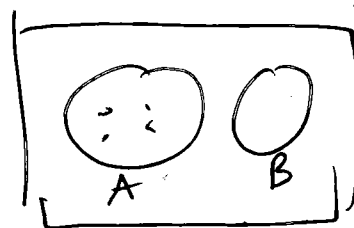
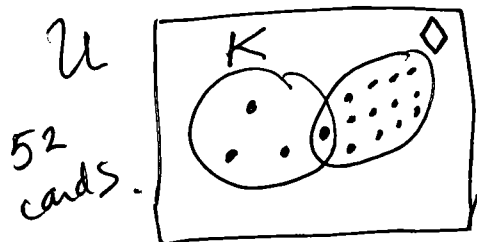
$$P(\text{Black} | \spadesuit) = 1$$

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

$$P(\text{Black}) = \frac{1}{2}$$

Can't throw out conditioning events.

Two independent events illustrated with a Venn diagram:



↑ not useful

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

$P(A) \neq 0$ not independent

$$P(A|B) = 0$$

DEFINITION Two events A and B are **independent** if $P(A|B) = P(A)$

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

Modified multiplication rule for $P(A \cap B)$ when the events are independent:

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

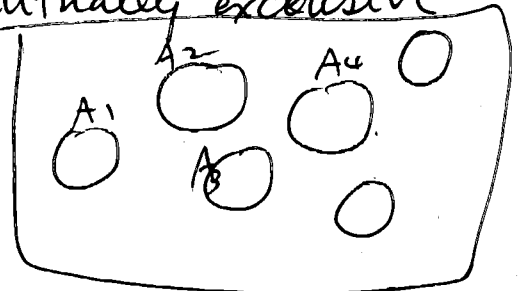
$$= \underbrace{P(B|A)} \cdot P(A) = P(B) \cdot P(A)$$

Generalize to k mutually independent events:

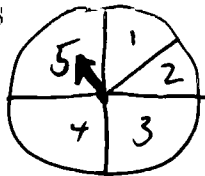
A_1, A_2, \dots, A_k Mutually ind^cf mutually exclusive

then $P(A_1 \cap A_2 \cap \dots \cap A_k)$
 $= P\left(\bigcap_{i=1}^k A_i\right) = \prod_{i=1}^k P(A_i)$

EXAMPLE Independent or dependent?



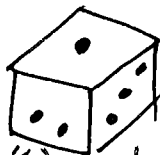
1. Spinners



results from different spins are independent

not mutually independent.

2. Dice

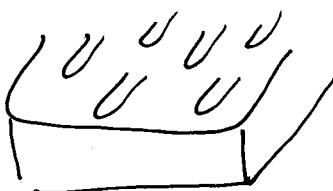


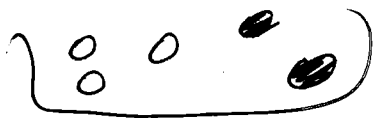
Rolling dice or roll the same die many times

$P("6" \cap "6") = \frac{1}{6} \cdot \frac{1}{6}$ results are independent.

3. Decks of cards, or drawers with different coloured socks

draw cards from a deck
 results are $\begin{matrix} \text{w R} \\ \text{w/o R} \end{matrix}$ independent
 dependent





4. Pots with different numbers: results from repeated draws with replacement are independent

vs. without replacement. dependent.

EXAMPLE If the events M: "being a male" and F: "failing the course" are independent, and you know that in a class of 84, 72 passed, and 49 are female. Fill in a two-by-two table classified by male, female, pass, and fail.

	Male	Female	
Pass	30	42	72
Fail	5	7	12
	35	49	84

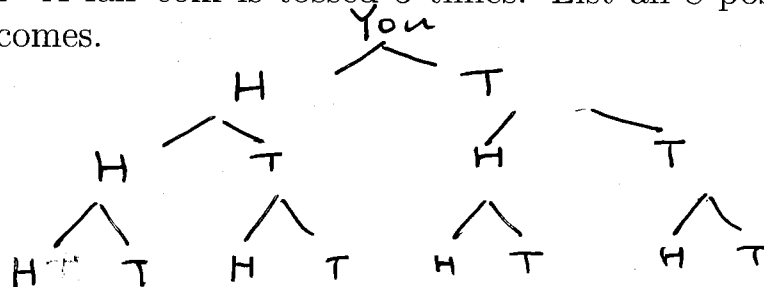
$P(M \cap F) = P(M) \cdot P(F)$
 $\frac{42}{84} = \frac{30}{84} \times \frac{12}{84}$
 $\frac{42}{84} = \frac{35}{84} \times \frac{12}{84}$

Are the events "being female": F_e and "Pass": P independent?

$$P(F_e \cap P) = \frac{42}{84} \stackrel{?}{=} \frac{49}{84} \times \frac{72}{84}$$

Yes, independent events.

EXAMPLE A fair coin is tossed 3 times. List all 8 possible equally likely outcomes.



1st time
2nd time

If we define the following events,

1. A: a head on each of the first 2 tosses
2. B: a tail on the 3rd toss

$$A = \{HH, HT, TH, TT\}$$

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

$$B = \{HTT, THT, TT\}$$

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

3. C: exactly 2 tails in the 3 tosses, $C: \{HTT, THT, TTH\}$

Determine whether events are pairwise independent.

$$P(C) = \frac{3}{8}$$

$$P(A \cap B) = \frac{1}{8} \quad \text{so independent } A, B. \quad P(A) \cdot P(B) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}$$

$$P(A \cap C) = 0, \text{ not independent } A, C. \quad P(A) \cdot P(C) = \frac{1}{4} \times \frac{3}{8} = \frac{3}{32}$$

$$P(B \cap C) = \frac{2}{8} = \frac{1}{4} \neq P(B) \cdot P(C) = \frac{1}{2} \times \frac{3}{8} = \frac{3}{16} \quad \text{not independent } B, C.$$

Reliability of systems *Reliability is $P(\text{system works})$*

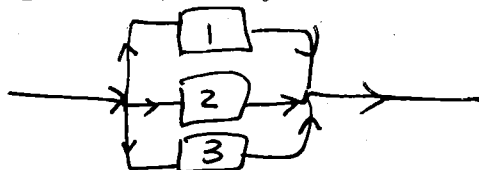
Electrical components or communication networks arranged

1. in **series**, the system works only if *all components $\boxed{1}$, $\boxed{2}$, $\boxed{3}$ work.*



the system fails if *at least one fails.*

2. in **parallel**, the system works if *at least one of $\boxed{1}$, $\boxed{2}$, $\boxed{3}$ works*



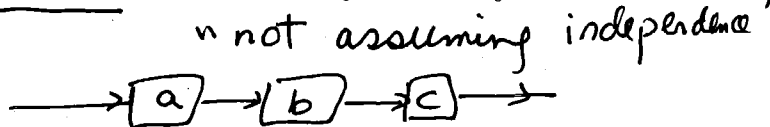
the system fails only if *all components fail.*

Need to find $P(\text{system works})$ and $P(\text{system fails})$ "assuming independent components." *simplified.*

$A \cup B$, $A \cap B$
 event A or event B event A and event B.

EXAMPLE Let A be the event that component a fails; B, and C likewise defined for components b and c respectively. Suppose $P(A) = 0.05$, $P(B) = 0.02$, $P(C) = 0.03$, estimate the reliability of the system a, b, c connected in series.

We know that the system fails if
 So, $P(\text{system fails}) =$



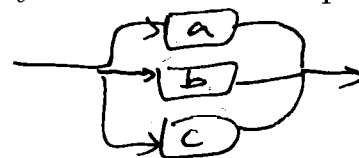
$$P(A \cup B \cup C) \leq P(A) + P(B) + P(C) \\ = 0.05 + 0.02 + 0.03 = 0.1$$

So, $P(\text{system works}) =$

$$= 1 - P(\text{system fails}) \\ \geq 1 - 0.1 = 0.9$$

Similarly, estimate the probability that the system works if components a, b, and c are connected in parallel.

$P(\text{system works})$



$$= P(A' \cup B' \cup C')$$

De Morgan's Law

$$= P((A \cap B \cap C)')$$

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

$$= 1 - P(A \cap B \cap C) \geq 1 - \min\{P(A), P(B), P(C)\} \\ = 1 - 0.02 = 0.98$$

When we know that the components are independent, we can compute exact reliability calculations.

Let us compute exactly the reliability of the systems above:

1) in series

$$P(\text{system works}) = P(A' \cap B' \cap C')$$

$$= P(A') \cdot P(B') \cdot P(C')$$

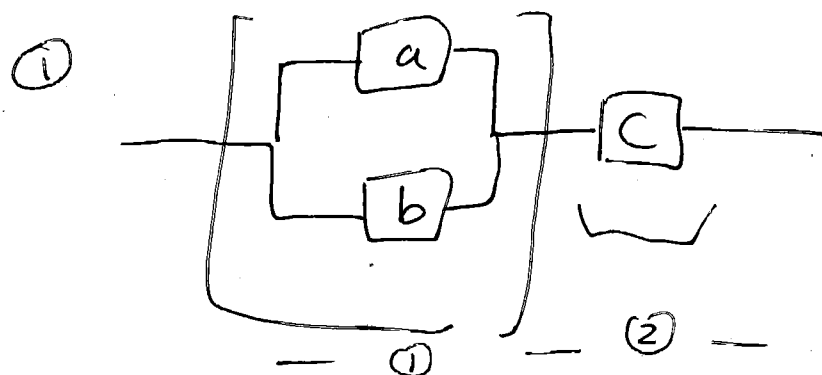
$$= (1 - 0.05) \cdot (1 - 0.02) \cdot (1 - 0.03)$$

$$\stackrel{54}{=} 0.95 \times 0.98 \times 0.97 \\ \approx 0.90307$$

2) in Parallel

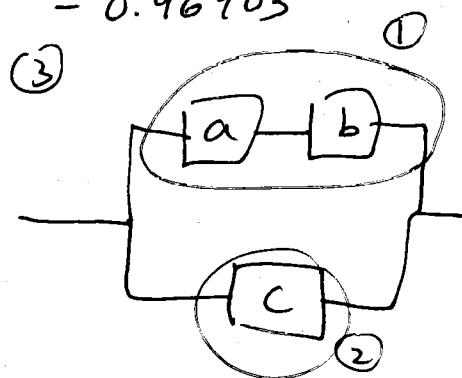
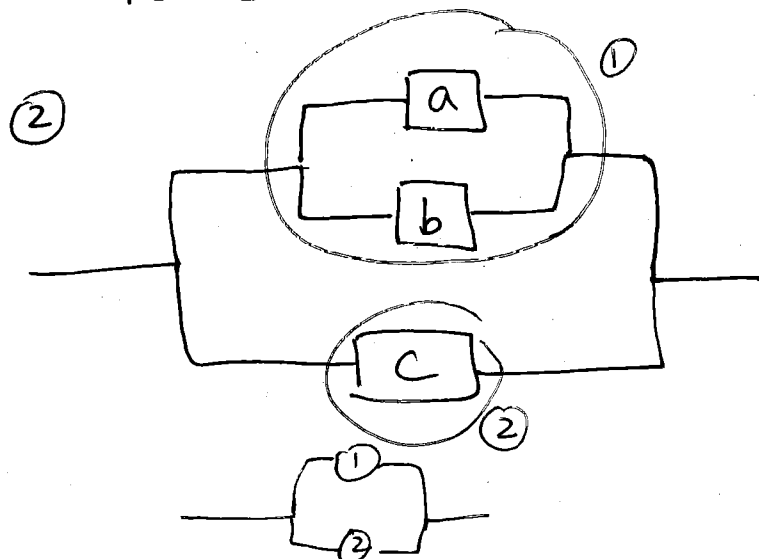
$$\begin{aligned}
 P(\text{system works}) &= 1 - P(A \cap B \cap C) \\
 &= 1 - P(A) \cdot P(B) \cdot P(C) \\
 &= 1 - 0.05 \times 0.02 \times 0.03 \\
 &= 0.99997
 \end{aligned}$$

EXAMPLE Let us consider a combined construction and compute its reliability, ie, $P(\text{system works})$.



0.96903 series

$$\begin{aligned}
 P(①) &= 1 - 0.05 \times 0.02 \\
 &= 0.999 \\
 P(②) &= 1 - 0.03 \\
 &= 0.97 \\
 P(① \cap ②) &= 0.999 \times 0.97 \\
 &= 0.96903
 \end{aligned}$$



Practice exercises in Section 2.5: 71, 73, 77, 79, 83, 85, 87.

Hand in § 2.4 & 2.5 Text
55 + 4th hr.

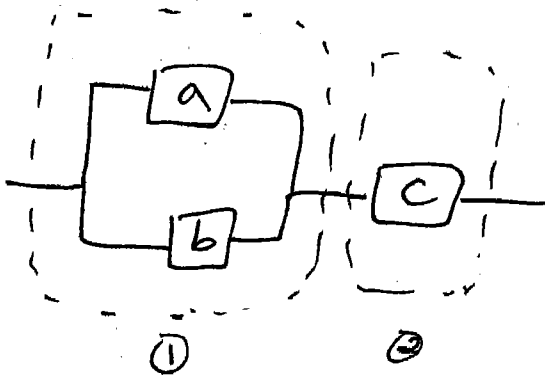
When independence assumed,
 P. sys series

$$a) P(A' \cap B' \cap C') = (1 - P(A))(1 - P(B)) \cdot (1 - P(C))$$

b) Parallel.

$$\begin{aligned} &P(\text{system works}) \\ &= 1 - P(A \cap B \cap C) \\ &= 1 - P(A) \cdot P(B) \cdot P(C) \end{aligned}$$

eg.



in series.

$$\textcircled{1} P(\textcircled{1} \text{ works})$$

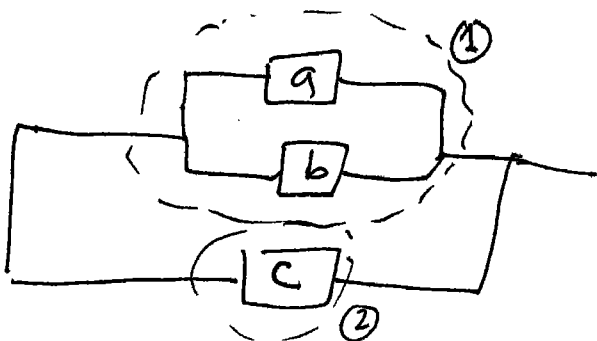
$$= 1 - P(A \cap B)$$

$$= 1 - P(A) \cdot P(B)$$

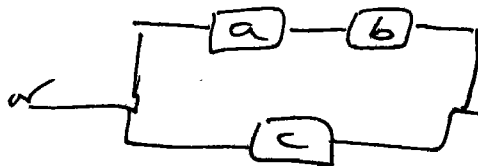
$$\textcircled{2} 1 - P(C) = P(\textcircled{2} \text{ works})$$

$$P(\textcircled{1} \cap \textcircled{2}) = (1 - P(A) \cdot P(B)) \cdot (1 - P(C))$$

both work



① ② in parallel.



Chapter 3: Discrete Random Variables and Probability Distributions

3.1 Random Variables

DEFINITION For a given sample space S of some experiment, a **random variable (rv)** is any rule that associates a number with each outcome in S . In mathematical language, a random variable is a function whose domain is the sample space S and whose range is the set of reals \mathbb{R} .

EXAMPLE When you register for a course, you will either get into the course, or be on the waitlist, or be rejected. With $S = \{I, W, R\}$, define a random variable X by

$$X(I) = 1, \quad X(W) = 0.5, \quad X(R) = 0.$$

DEFINITION Any random variable whose only possible values are 0 and 1 is called a **Bernoulli random variable**.

$$\text{Range}(X) = \{0, 1\}$$

EXAMPLE Starting at a fixed time, each car entering an intersection is observed to see whether it turns left (L), right (R), or goes straight ahead (A). The experiment terminates as soon as a car is observed to turn left. Let X be the number of cars observed. What are the possible X values? List five outcomes and their associated X values.

$$\text{range}(X) = \{1, 2, \dots\} = \mathbb{N}$$

Possible outcomes in domain of X : L, RL, AL
 $X(RRARL) = 5$

DEFINITION A discrete random variable is a r.v. whose possible values, i.e. range, either consist of a finite set (eg $\{0, 2, 5, 7\}$) or can be listed in a infinite sequence⁵⁷. ("countably" infinite sequence).

A **continuous** random variable requires both of the following to apply:

1. Interval condition: all the numbers in a single interval or union of disjoint intervals.

e.g. Range = $[0, 1]$

4 Range = $(-3, -2] \cup [1, 5)$

2. Point probability: no possible value of the random value has a + prob.

i.e. $P(X=c) = 0$ for all $c \in S$

EXAMPLE 1. Select 2 lights from a box of 7 working and 3 defective lights. Define the rv, W to be the number of working lights.

$W(L_1, L_2) =$

$W(D, D) = 0$

finite set

Range(W) = $\{0, 1, 2\}$

$P(0) =$

$P(2) =$

$\begin{matrix} W & W & W & 7W \\ W & D & W & \\ W & D & W & D \end{matrix} \begin{matrix} 3D \end{matrix}$

$W(\underline{D}, \underline{W}) = 1$
 $\in S$

2 \ 1	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

2. Roll 2 fair dice. Define the following rv's: T is the total of the dice; M is the maximum value of the dice; D is the absolute difference of the dice.

Range(T) = $\{2, 3, 4, \dots, 12\}$, Range(M) = $\{1, 2, 3, \dots, 6\}$, Range(D) = $\{0, 1, 2, \dots, 5\}$

3. Flip a coin until a head appears. Let X be the number of flips needed before the first HEAD.

Range(X) = $\{0, 1, 2, 3, \dots\}$ } = ∞
TH TTTH TTTH = ∞

$X(H) = 0$

The previous 3 examples have been discrete
r.v.

4. Randomly select a transistor from a production run. Let X be the lifetime of the transistor.

$$X(\text{Transistor } a) = a \in \mathbb{R}$$

continuous
random v

$$[0, \infty)$$

Practice exercises in Section 3.1: All odds from 1 to 9.

3.2 Probability Distributions for Discrete Random Variables

EXAMPLE Capilano University's Pure and Applied Science Division has a lab with six computers reserved for students taking statistics. Let X denote the number of these computers that are in use at a particular time of day. Suppose that the probability distribution of X is as given in Table 3.1.

x	0	1	2	3	4	5	6	
$p(x)$	0.05	0.10	0.15	0.25	0.20	0.15	0.10	$\sum p(x) = 1$

Table 3.1: Computer use

1. Find the probability that at least one computer is in use.

$$\begin{aligned}
 P(X \geq 1) &= p(1) + p(2) + p(3) + p(4) + p(5) + p(6) \\
 &= 1 - p(0) = 1 - 0.05 = 0.95
 \end{aligned}$$

2. The administration needs to cut the budget and would like to close the lab if less than half of the computers is in use. Find the probability.
 0, 1, 2,

$$\begin{aligned}
 P(X < 3) &= p(0) + p(1) + p(2) \\
 &= 0.30
 \end{aligned}$$

DEFINITION The **probability distribution** or **probability mass function** (pmf) of a discrete rv is defined for every number x by

$$p(x) = P(X = x) = P(\text{all } s \in S : X(s) = x).$$

discrete rv.
 \uparrow actual value X takes

Remember that for any pmf, we always have the conditions that

$$p(x) \geq 0 \quad \text{and} \quad \sum_x p(x) = 1$$

EXAMPLE Consider whether the next person buying a computer at Future Shop buys a laptop or a desktop model. Let

Best Buy

$$X = \begin{cases} 1, & \text{if the costumer buys a desktop,} \\ 0, & \text{if the costumer buys a laptop.} \end{cases}$$

If 20% of all purchasers during that week select a desktop, the pmf for X is

$$p(0) = P(X = 0) = 0.8, \quad = 1 - 0.2$$

$$p(1) = P(X = 1) = 0.2,$$

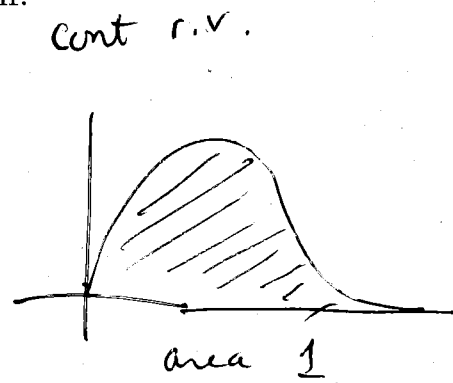
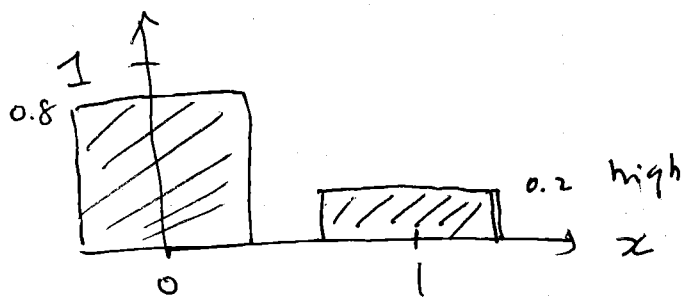
$$p(x) = P(X = x) = 0, \quad \text{if } x \neq 0 \text{ and } x \neq 1.$$

Equivalently,

$$p(x) = \begin{cases} 0.8, & \text{if } x = 0, \\ 0.2, & \text{if } x = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 3.1 on the next page is a picture of this pmf, called a **line graph**. The variable X is, of course, a Bernoulli rv and $p(x)$ is a Bernoulli pmf.

Line graph vs. probability histogram for the pmf.



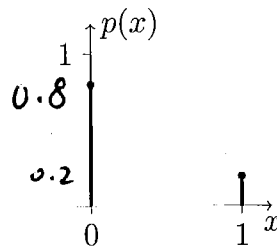


Figure 3.1: The line graph for $p(x)$

A parameter of a probability distribution

DEFINITION Suppose $p(x)$ depends on a quantity that can be assigned any one of a number of possible values, with each different value determining a different probability distribution. Such a quantity is called a **parameter** of the distribution. The collection of all probability distributions for different values of the parameter is called a **family** of probability distributions.

α : parameter

EXAMPLE Consider web registration at Capilano University. Starting at a fixed time, we observe whether new registrants are foreign or domestic until we see a foreign student (F). Let $\alpha = P(F)$, assume that successive registrants are independent, and define the rv X by $x =$ number of registrants observed. Then

DDDF

$$p(1) = P(X = \underline{1}) = P(\underline{F}) = \alpha$$

$$p(2) = P(X = \underline{2}) = P(\underline{DF}) = \underbrace{P(D)}_s \cdot P(F) = (1 - \alpha)\alpha$$

and prob mass fn $p(x)$

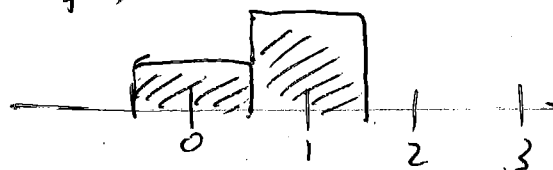
$$p(3) = P(X = 3) = P(DDF) = (1 - \alpha)^2 \alpha$$

Continuing in this way, a general formula emerges:

$$p(x) = \begin{cases} (1 - \alpha)^{x-1} \alpha & x = 1, 2, 3, \dots \\ 0, & \text{otherwise.} \end{cases} \quad (3.1)$$

The parameter α can assume any value between 0 and 1. Expression 3.1 describes the family of **geometric distributions**.

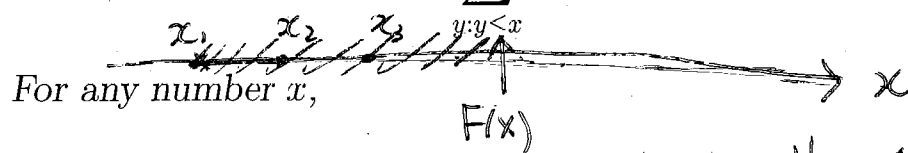
Recall pmf : $p(x)$



The cumulative distribution function

DEFINITION The **cumulative distribution function** (cdf) $F(x)$ of a discrete rv X with pmf $p(x)$ is defined for every number x by

$$F(x) = P(X \leq x) = \sum_{y: y \leq x} p(y)$$

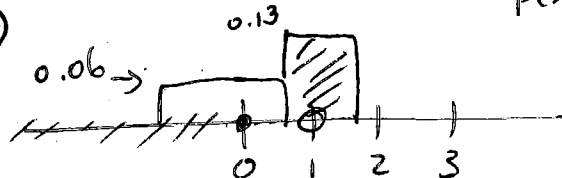


$F(x)$ is the prob that the observed value of X will be at most x .

EXAMPLE A consumer organization that evaluates new automobiles customarily reports the number of major defects in each car examined. Let X denote the number of major defects in a randomly selected car of a certain type. The cdf of X is as follows:

$$F(x) = \begin{cases} 0, & x < 0, \\ 0.06, & 0 \leq x < 1, \quad [0, 1) \\ 0.19, & 1 \leq x < 2, \quad [1, 2) \\ 0.39, & 2 \leq x < 3, \\ 0.67, & 3 \leq x < 4, \\ 0.92, & 4 \leq x < 5, \\ 0.97, & 5 \leq x < 6, \\ 1, & 6 \leq x. \end{cases}$$

X discrete r.v.
 $\{0, 1, 2, 3, \dots\}$
 $p(x)$



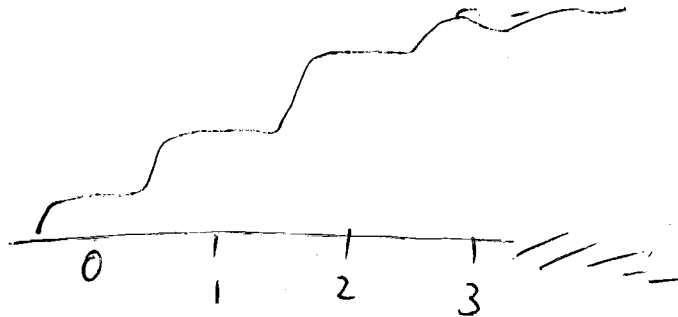
$$p(1) = 0.19 - 0.06$$

Calculate the following probabilities directly from the cdf:

1. $p(2)$, that is, $P(X = 2)$.

$$= F(2) - F(2-)$$

$$= 0.39 - 0.19 = 0.2$$



$$2. P(X > 3) =$$

$$p(4) + p(5) + p(6)$$

$$= F(6) - F(3) = 1 - 0.67 = 0.33$$

$$3. P(2 \leq X \leq 5),$$

$$= F(5) - F(2-)$$

$$= 0.97 - 0.19 = 0.78$$

$$4. P(2 < X < 5).$$

$$= F(4) - F(2)$$

$$= 0.92 - 0.39 = 0.53$$

PROPOSITION For any two numbers a and b with $a \leq b$,

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

where " $a-$ " means the largest possible X value
strictly less than a .

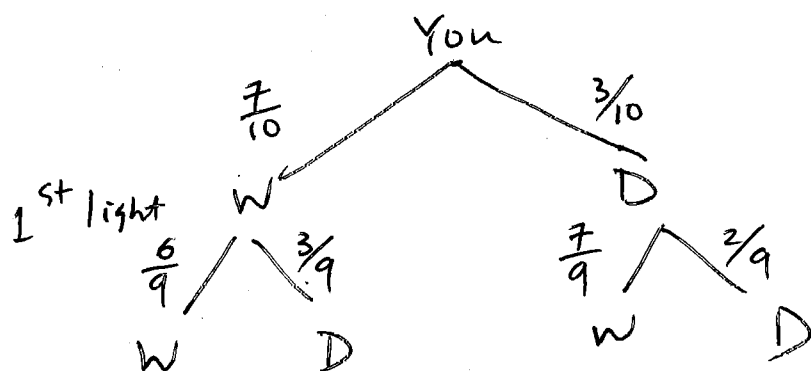
$$x < a$$

insert to p.65 LN.

65a)

7 W's
3 D's

r.v. X : # of working lights out of 2
taken without replacement.



$P(W, W)$

$$P(2) = \frac{7}{10} \times \frac{6}{9} = \frac{7}{15}$$

$$P(1) = \frac{7}{10} \times \frac{1}{3} + \frac{3}{10} \times \frac{7}{9}$$

$$= \frac{14}{30} = \frac{7}{15}$$

$$P(0) = \frac{3}{10} \times \frac{2}{9} = \frac{1}{15}$$

$$\sum f(x) = 1$$

EXAMPLE 1. Verify that $f : f(x) = \frac{x+2}{25}$ for $x = 1, 2, \dots, 5$ cannot serve as a pmf for a discrete rv.

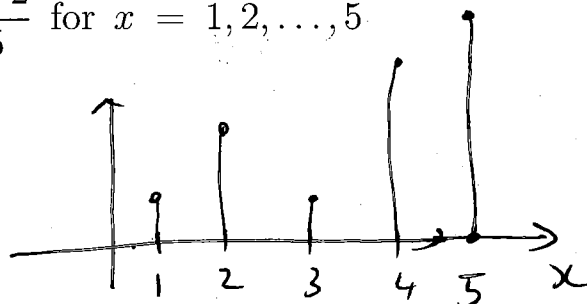
$$f(1) = \frac{3}{25}$$

$$f(4) = \frac{6}{25}$$

$$f(2) = \frac{4}{25}$$

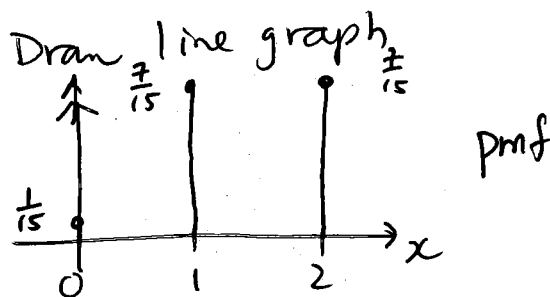
$$f(5) = \frac{7}{25}$$

$$f(3) = \frac{5}{25} = \frac{1}{5}$$



2. Revisit picking 2 working lights problem out of 7 working and 3 defective ones. Find its pmf.

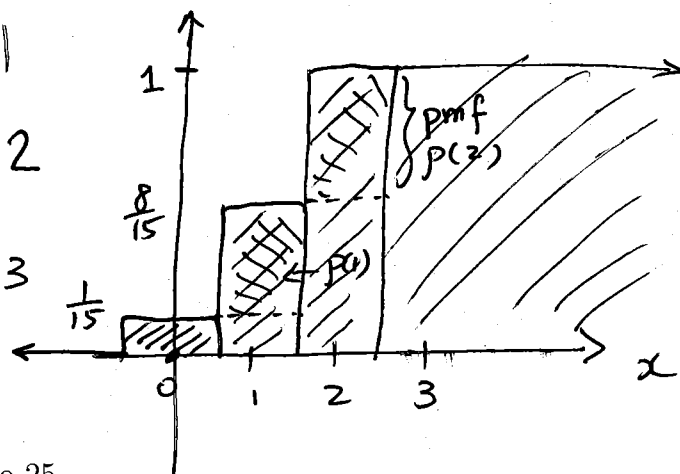
x	0	1	2
$p(x)$	$\frac{1}{15}$	$\frac{7}{15}$	$\frac{7}{15}$



3. Rewrite the previous pmf as a cdf and draw its corresponding step function.

cumulative distribution fn $F(x)$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{15} & 0 \leq x < 1 \\ \frac{8}{15} & 1 \leq x < 2 \\ 1 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$



Practice exercises in Section 3.2: all odds from 11 to 25.

3.3 Expected Values

Recall how you compute averages:

Average of n numbers x_1, x_2, \dots, x_n is

$$\frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

EXAMPLE Consider 1000 students in the science division and define the rv, X to be the number of courses taken for a randomly selected student. The following pmf is given.

	x	1	2	3	4	5	6	7	
pmf	$p(x)$	0.05	0.10	0.2	0.5	0.1	0.025	0.025	
$p(x) \times 1000$	students	50	100	200	500	100	25	25	Total ← 1000

Table 3.2: Science students and courses

On average, how many courses does a randomly selected student take?

$$\begin{aligned} \text{average} &= \frac{50 \times 1 + 100 \times 2 + 200 \times 3 + 500 \times 4 + 100 \times 5 + 25 \times 6 + 25 \times 7}{1000} \\ &= \frac{50}{1000} \times 1 + \frac{100}{1000} \times 2 + \frac{200}{1000} \times 3 + \frac{500}{1000} \times 4 + \frac{100}{1000} \times 5 + \frac{25}{1000} \times 6 + \frac{25}{1000} \times 7 \\ &= \sum_x p(x) \cdot x = \text{average \# of courses} \end{aligned}$$

DEFINITION Let X be a discrete rv with set of possible values D and pmf $p(x)$. The **expected value** or **mean value** of X , denoted by $E(X)$ or μ_X or just μ , is

" $m\mu$ "

$$E(X) = \mu_X = \sum_{x \in D} xp(x)$$

$$\left\{ \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right] \right\} \quad ((((()))))$$

EXAMPLE Suppose the number of plants of a particular type found in a rectangular sampling region (called quadrat by ecologists) in a certain geographic area is a rv X with pmf

$$p(x) = \begin{cases} c/x^3, & x = 1, 2, 3, \dots \\ 0, & \text{otherwise.} \end{cases}$$

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \dots = \frac{\pi^2}{6}$$

Is $E(X)$ finite?

$$E(X) = \sum_x p(x) \cdot x = \sum_{x=1}^{\infty} \frac{c}{x^3} \cdot x = \sum_{x=1}^{\infty} \frac{c}{x^2} = c \cdot \frac{\pi^2}{6}$$

$E(X)$ is the average # of plants in a quadrat.

DEFINITION If the rv X has a set of possible values D and pmf $p(x)$, then the expected value of any function $h(x)$, denoted by

is computed by $E[h(X)]$, $\mu_{h(X)}$, μ_h

$$E[h(X)] = \sum_D h(x)p(x)$$

EXAMPLE Let X be the outcome when a fair die is rolled once. If before the die is rolled you are offered either $\frac{1}{3.5}$ dollars or $\left(\frac{1}{X}\right)$ dollars, would you accept the guaranteed amount or would you gamble?



$$E\left(\frac{1}{X}\right) = \sum_{x=1}^6 \frac{1}{x} \cdot \frac{1}{6} = \frac{1}{6} \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} \right) = \frac{1}{6} \left(\frac{60+30+20+15+12+10}{60} \right)$$

$$\text{c.f. } \frac{1}{3.5} \approx 0.2857$$

Compare with

Note that in general $\boxed{\frac{1}{E(X)} \neq E(1/X)}$

$$E(X) = \sum x \cdot \frac{1}{6} \quad 67$$

$$= \frac{1}{6} (1+2+3+4+5+6) = \frac{21}{6} = 3.5$$

$$= \frac{1}{6} \cdot \frac{147}{60} = \frac{147}{360}$$

$$\approx 0.408\bar{3}$$

Bet for $\frac{1}{X}$

$E()$ is not multiplicative

$$\frac{d}{dx} (a f(x) + b) = a \frac{df(x)}{dx} + 0$$

$$\int a f(x) + b \, dx = a \cdot \int f(x) \, dx + \int b \, dx$$

Linearity of expected value, variance, σ^2 , and rules of variance

PROPOSITION Linearity of E .

$$E(aX + b) = a E(X) + b \quad \begin{array}{l} a, b \text{ constants} \\ \text{linear comb} \end{array}$$

$$E(X), \mu_X$$

DEFINITION Let X have pmf $p(x)$ and expected value μ . Then the **variance** of X , denoted by $V(X)$ or σ_x^2 , or just σ^2 , is

$$V(X) = \sum_D \underbrace{(x - \mu)^2 \cdot p(x)} = E((X - \mu)^2)$$

The **standard deviation** (SD) of X is $\sigma_X = \sqrt{\sigma_x^2} = \sqrt{V(X)}$

$$\sigma_{X+b} = \sigma_X \quad V(X) = E(X^2) - (E(X))^2, \quad V(aX+b) = a^2 \sigma_X^2, \quad \sigma_{aX+b} = |a| \sigma_X$$

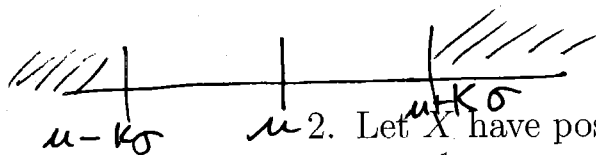
EXAMPLE A result called **Chebyshev's inequality** states that for any probability distribution of a rv X and any number k that is at least 1, $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$. In words, the probability that the value of X lies at least k standard deviations from its mean is at most $1/k^2$.

1. What is the value of the upper bound for $k = 2$? $k = 3$? $k = 4$?

$$k = 5? \quad k = 10? \quad \frac{1}{100}$$

$$\frac{1}{4}, \quad \frac{1}{9}, \quad \frac{1}{16}$$

$$|X - \mu| \leq \frac{1}{\sqrt{5}} \leq \frac{1}{k^2}$$



2. Let X have possible values $-1, 0$, and 1 , with probabilities $\frac{1}{18}, \frac{8}{9}$, and $\frac{1}{18}$, respectively. What is $P(|X - \mu| \geq 3\sigma)$, and how does it compare to the corresponding upper bound?

$$\begin{array}{c|ccc} X & -1 & 0 & 1 \\ \hline p(x) & \frac{1}{18} & \frac{8}{9} & \frac{1}{18} \end{array}, \quad \mu = \sum x p(x) = (-1) \cdot \frac{1}{18} + 0 \cdot \frac{8}{9} + 1 \cdot \frac{1}{18} = 0$$

$$\sigma = \sqrt{\sum_{x \in \{-1, 0, 1\}} (x - 0)^2 p(x)} = \sqrt{1 \cdot \frac{1}{18} + 0 \cdot \frac{8}{9} + 1 \cdot \frac{1}{18}}$$

Two computational short cuts from two propositions in Section 3.3.

$$P(|X - 0| \geq 3 \cdot \frac{1}{3}) \leq \frac{1}{k^2} = \frac{1}{3^2} = \frac{1}{9} \quad \sqrt{\frac{1}{9}} = \frac{1}{3}$$

$$= P(|X| \geq 1) = \frac{1}{18} + \frac{1}{18} = \frac{2}{18} = \frac{1}{9}$$

$f_{X=-1} \quad X=1$

X : demand.
 $\#$ of cheesecakes sold on a given day

EXAMPLE Bess, a bakery manager, knows that the number of cheesecakes she can sell on a given day (*demand* for cheesecakes) is a random variable X with a pmf $p(x) = \frac{1}{6}$, $x = 0, 1, \dots, 5$. Suppose there is a profit of \$2.5 for each cheesecake sold, but a loss of \$1.00 for each cheesecake remaining unsold. Find the expected net profit for a day on which she bakes 4 cheese cakes.

Let $N(x)$ be the net profit as a function of x . Find $N(x)$. $h(x)$

x	0	1	2	3	4	5
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$N(x)$	-4	-0.5	3	6.5	10	10

$\sum p(x) = 1$ $N(x) = \begin{cases} 2.5x & x=0,1,2,3,4 \\ -1 \cdot (4-x) & 5 \end{cases}$

Make a table of values for x , $p(x)$, $N(x)$, and $N(x)p(x)$ before you compute $E(N(X))$.

$$E(N(X)) = \sum N(x) \cdot p(x)$$

$$= \sum_x N(x) \cdot \frac{1}{6}$$

$$= \frac{1}{6} (-4 - 0.5 + 3 + 6.5 + 10 + 10)$$

$$= \frac{25}{6} \approx 4.1\bar{6}$$

Expected net profit is \$4.17

Practice exercises in Section 3.3: odds from 29 to 41.

Recall Geometric dist.

$$P(x) = \begin{cases} \alpha^{x-1} \cdot (1-\alpha), & \alpha \\ 0 & \text{other values} \end{cases}$$

$\alpha = \text{Prob of } S$

3.4 The Binomial Probability Distribution

The set-up: When experiments conform either exactly or approximately to the following list of requirements:

1. Fix n in advance of the experiment. The experiment consists of a sequence of n smaller experiments called **trials**. n positions

2. Each trial can result in one of the same two possible outcomes:

S : Success

F : Failure

3. The trials are independent,

4. The probability of success $P(S)$ is constant from trial to trial; denote it by p .

$$P(F) = 1 - p = q$$

DEFINITION An experiment for which all four conditions are satisfied is called a **binomial experiment**

EXAMPLE 1. coin tossing, thumb tag tossing, knife throwing, ...

2 outcomes, fix 50 trials, independent.

$$P(H) = \frac{1}{2}, \quad P(\text{up}) = \text{constant}$$

2. boy or girl? Observe 50 pregnancies.

3. domestic or foreign worker? need a fixed # of workers to study

4. Sample 100 CapU students and record whether the student has type AB blood.

39 S, 10 F

L

✓ $\frac{S}{50th}$

$n = 29$

5. For 28 days in February this year, record whether or not it rains that day in North Vancouver. R, N.

not independent.

EXAMPLE Let us consider two examples of the same nature but different population, small vs. large.

ex. Small village 50 households, 40 have insurance, 10 don't

Success: the event that a household has insurance

1 2 3 4 5 50

$$P(1^{st} \text{ house } S) = \frac{40}{50} = 0.8$$

$$P(2^{nd} S) = P(2^{nd} S | 1^{st} S) + P(2^{nd} S | 1^{st} F)$$

$$= \frac{40}{50} \cdot \frac{39}{49} + \frac{10}{50} \cdot \frac{40}{49}$$

$$P(n^{th} \text{ house } S) = \frac{40(10+39)}{50 \cdot 49} = \frac{4}{5} = 0.8$$

Summarize:

Consider sampling without replacement from a dichotomous population of size N.

Sample size 5% of N

However

$$P(S \text{ on } 5^{th} \text{ trial} | SSSS) = \text{different}$$

$$P(S \text{ on } 5^{th} \text{ trial} | FFFF) = \text{values.}$$

72

Don't have independence in general)

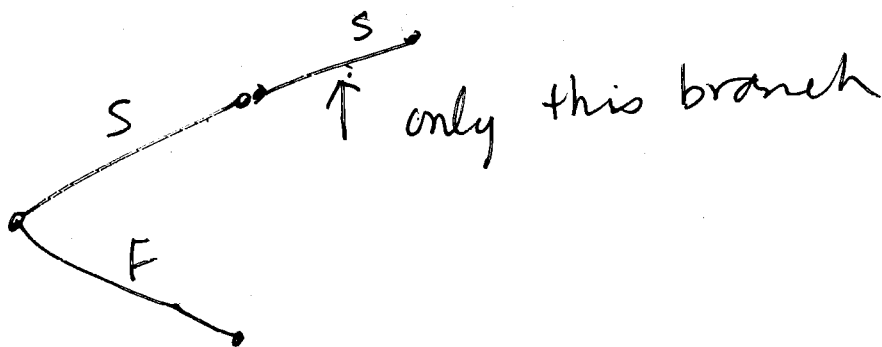
Conditional prob

500,000 households

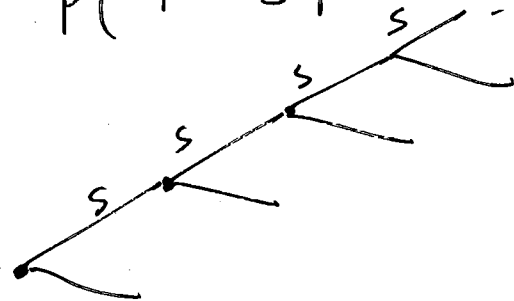
400,000 insured

S: house insured

$$P(2^{\text{nd}} S \mid 1^{\text{st}} S) = \frac{399,999}{499,999}$$



$$P(10^{\text{th}} S \mid \text{first } 9 \text{ S's}) = \text{branch}$$



$$= \frac{399,991}{499,991}$$

replace 72 a)

DEFINITION The **binomial random variable** X associated with a binomial experiment consisting of n trials is defined as

$$X = \# \text{ of } S\text{'s among } n \text{ trials}$$

$S : \text{success}$

We use $b(x; n, p)$ to denote pmf of a binomial rv X on two parameters n and p .
 $P(S) = p$

EXAMPLE Flipping a coin 50 times is a binomial experiment. However, not all rv's defined are binomial rv's. Determine which of the following is a binomial rv.

1. Let X be the number of flips until (before) the first HEAD appears. coin toss

List possible values: until H, TH, TTH, ... $\{1, 2, 3, \dots, 50\}$
before

2. Let Y be the number of TAILS observed in 50 trials. $\{0, 1, 2, 3, \dots, 49\}$

List possible values: $\{0, 1, 2, \dots, 50\}$

3. Let Z be the length of the longest run of HEADs. HHHH...H

List possible values: $\{0, 1, 2, \dots, 50\}$

4. Let W be the difference in the number of H's and T's. THTHTH

List possible values: $\{0, 2, 4, \dots, 48, 50\}$

Y is the only binomial r.v.

$$(S + F)(S + F)(S + F) = (S + F)^3$$

Let us study the outcomes and probabilities for a binomial experiment with three trials:

outcome	x	prob	$p(S) = p$	$q = 1 - p$
SSS	3	$\binom{3}{3} p^3$		
$\left\{ \begin{array}{l} SSF \\ SFS \\ FSS \end{array} \right\}$	2	$\binom{3}{2} p^2 q$		
$\left\{ \begin{array}{l} SFF \\ FSF \\ FFS \end{array} \right\}$	1	$\binom{3}{1} p q^2$		
FFF	0	$\binom{3}{0} q^3$		

THEOREM

$$b(x; n, p) = \begin{cases} \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} & x=0, 1, \dots, n \\ 0 & \text{otherwise} \end{cases}$$

Pascal's Δ

Notation in TI requires binompdf(n, p, x) entered in that order.

Let $X \sim \text{Bin}(n, p)$ denote that the random variable, X is a binomial random variable with parameters n and p where n means number of trials and p means $P(S)$ n trials.

PROPOSITION (MEAN, VARIANCE, AND STANDARD DEVIATION)

If $X \sim \text{Bin}(n, p)$, then

$$E(X) = n \cdot p, \quad V(X) = npq = np(1-p)$$

$$\sigma_x = \sqrt{npq}$$

You need to prove this on your own using the binomial expansion formula. See 4th hr activity.

EXAMPLE Application: Acceptance Testing

Before shipping a large batch of components, a manufacturer tests the batch to determine if it is acceptable. Because of the expense and/or nature of the tests, only a small number of components in each batch are tested to see if they are defective or not.

Suppose a batch is deemed acceptable if its proportion of defective components does not exceed 0.10 ($p \leq 0.10$).

The manufacturer decides to use the following test procedure:

- Randomly sample 10 (sample size) components and test them.
- Accept the batch if the number of defective items is 2 or less; otherwise, reject the batch. (Decision Rule)

Two types of errors can be made:

1. Reject the batch, even if the true proportion of defectives does not exceed 0.10.

What is the cost of this type of error?

- loss of revenue
- more testing
- resource cost, time, ...

2. Accept a bad batch; that is, skip a batch which has a true proportion of defectives in excess of 0.10. (Say 0.20).

What is the cost of this type of error?

- warranty repair cost
- reputation, loss customer base
- risk of bad components
- law suits.

Find the probability of making each of the above types of errors.

$$\text{binomcdf}(n, p, x)$$

S : the event of picking a defective component.

1. Probability of rejecting a good batch: If $p = 0.10$ and $n = 10$, we know that $X \sim \text{Bin}(10, 0.1)$, so

$$\begin{aligned} P(X > 2) &= 1 - \underbrace{P(X \leq 2)}_{\text{cdf}} \\ &= 1 - \text{binomcdf}(10, 0.1, 2) \\ &\approx \underbrace{0.07019}_{\text{reject if } P(X > 2)} \end{aligned}$$

$$X \leq 2$$

$$0.02000$$

2. Probability of passing a bad batch (with say $p = 0.20$): Now, $X \sim \text{Bin}(10, 0.2)$, so

$$P(X \leq 2) = \text{binomcdf}(10, 0.2, 2) \approx 0.6778$$

$$B(x; n, p)$$

Suppose we want to reduce this probability by changing the decision rule to: accept the batch only if one or fewer tested is defective. Recompute parts one and two and compare.

$$\begin{aligned} 1. P(\text{reject good}) &= 1 - P(X \leq 1) \\ p=0.1 \quad P(X > 1) &= 1 - \text{binomcdf}(10, 0.1, 1) \\ &\approx 0.2639 \uparrow \text{ from } 0.07019 \end{aligned}$$

$$\begin{aligned} 2. P(\text{accept bad}) &= P(X \leq 1) = \text{binomcdf}(10, 0.2, 1) \\ p=0.2. &\approx 0.3758 \downarrow \text{ from } 0.6778 \end{aligned}$$

Can we reduce the probabilities of both types of errors?

$n=10$, 20 Let us change the sample size in the text procedure to $n=20$. That is,

you accept a batch only if the number of defective items is 4 or less.

Recompute both probabilities.

$$\begin{aligned} 1. P(X > 4) &= 1 - P(X \leq 4) \\ &= 1 - \text{binomcdf}(20, 0.1, 4) \\ &\approx 0.04317 \end{aligned}$$

$$\begin{aligned} 2. P(X \leq 4) &\approx B(4; 20, 0.2) \approx 0.6296 \\ p=0.2 &\text{binomcdf}(20, 0.2, 4) \end{aligned}$$

accept
2 or fewer

Naturally, $P(A)$ where A is the event of accepting a bad batch depends on the actual proportion of bad components in the batch. Calculating $P(A)$ for different values of p and graphing them yields the **operating characteristic** (OC) curve of an acceptance sampling plan.

Practice exercises in Section 3.4: 47, 49, 55, 57, 59, 63, 65.

3.5 Hypergeometric and negative binomial distributions

Recall the binomial distribution: pmf of a binom dist

$$b(x; n, p) = \begin{cases} \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}, & x=0, 1, \dots, n \\ 0 & \text{otherwise.} \end{cases}$$

Conditions ① Fix n trials. ② dichotomous, S, F

③ $p(S) = \text{constant}$

Instead of the approximate probability model for sampling without replacement from a finite dichotomous population with a small sample size

n relative to population size N , the **hypergeometric distribution** is the exact probability model for the number of S 's in the sample. h .

The binomial rv X is the number of S 's when the number n of trials is fixed, whereas the **negative binomial distribution** arises from fixing the number of S 's desired and letting the number of trials be random.

Assumptions for the hypergeometric distribution: nb

1. A finite (size N) population where samples are taken.

2. Dichotomous, and there are M successes in the population. $\{S, F\}$ $\text{total } S\text{'s known among } N$

3. A sample of n individuals is selected without replacement in such a way that each subset of size n is equally likely to be chosen.

random
simple
sample

The random variable of interest is X the number of S 's in the sample. h n trials

The probability distribution of X depends on the parameters n , M , and N , so we wish to obtain $P(X = x) = h(x; n, M, N)$. M S 's N popul.

PROPOSITION If X is the number of S 's in a completely random sample of size n drawn from a population consisting of M S 's and

$X = \#$ of successes, population size of N then the probability distribution of X , called the **hypergeometric distribution**, is given by $\text{total } S$

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

\uparrow \uparrow \uparrow
 $\#$ of S 's Population n trials
 among n trials

$$\underbrace{\frac{S}{1} \frac{S}{2} \frac{S}{3} \dots \frac{S}{M}}_{M's} \underbrace{\frac{F}{M+1} \frac{F}{M+2} \dots \frac{F}{N}}_{N-M \text{ F's}} \quad \text{choose } n \text{ positions}$$

for x , an integer, satisfying

$$\text{what value of } x: \max(0, n-N+M) \leq x \leq \min(n, M)$$

Because if $n < M$, then largest $X = n$
if $M < n$, then largest $X = M$

if $N-M \geq n$, $\min X = 0$

$$N-M < n, \min X = n - (N-M) = n - N + M > 0$$

PROPOSITION (MEAN AND VARIANCE) The mean and variance of the hypergeometric rv X having pmf $h(x; n, M, N)$ are

$$E(X) = n \cdot \underbrace{\frac{M}{N}}_{p \text{ from binom.}}, \quad V(X) = \underbrace{\left(\frac{N-n}{N-1}\right)}_{\text{recall binom.}} n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$$

recall binom.
 $X \sim \text{Bin}(n, p)$
 $E(X) = np$
 $V(X) = npq$

The finite population correction factor is $\frac{N-n}{N-1}$

EXAMPLE An instructor who taught two sections of engineering statistics last term, the first with 20 students and the second with 30, decided to assign a term project. After all the projects had been turned in, the instructor randomly ordered them before grading. Consider the first 15 graded projects. Let S denote the event that the proj is from section 2.

1. What is the probability that exactly 10 of these are from the second section?

Is this hypergeometric, X : # of projects among 15 from section 2.
 $n = 15$, $M = 30$, $N = 50$
 $x = 10$

$$h(x; n, M, N) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

$$n - N + M = 15 - 50 + 30 < 0 \quad (0 \leq x \leq 15) \quad h(10; 15, 30, 50) = \frac{\binom{30}{10} \binom{20}{5}}{\binom{50}{15}}$$

2. What is the probability that at least 10 of these are from the second section?

$$\begin{aligned} & h(10; 15, 30, 50) \\ & + h(11; 15, 30, 50) \\ & + h(12; \quad \quad \quad) \\ & + h(13; \quad \quad \quad) \\ & + h(14; \quad \quad \quad) \\ & + h(15; \quad \quad \quad) = P_2 \end{aligned}$$

3. What is the probability that at least 10 of these are from the

same section $P_1 (X \geq 10) = h(10; 15, 20, 50)$

$$\begin{aligned} & h(11; 15, \quad \quad) \\ & h(12; 15, \quad \quad) \\ & h(13; \quad \quad \quad) \\ & h(14; \quad \quad \quad) \\ & h(15; \quad \quad \quad) \end{aligned}$$

$$P_1 + P_2$$

4. What are the mean value and standard deviation of the number among these 15 that are from the second section?

$$\begin{aligned} E(X) &= n \cdot \frac{M}{N} \\ &= 15 \cdot \frac{30}{50} \\ &= 9 \end{aligned}$$

$$\begin{aligned} \sigma_x &= \sqrt{V(X)} = \sqrt{\left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)} \\ &= \sqrt{\left(\frac{50-15}{50-1}\right) \cdot 15 \cdot \frac{30}{50} \cdot \frac{20}{50}} \end{aligned}$$

5. What are the mean value and standard deviation of the number of projects not among these first 15 that are from the second section?

the rest $n = 35$.

- ① geometric dist.
- ② binomial dist.
- ③ hypergeometric dist
- ④

pmf. pdf — cont.
X:

→ Negative binomial distribution

Experimental conditions for the **negative binomial** rv and distribution:

binom
dist.

1. The experiment consists of a sequence of independent trials.
2. Dichotomous results.
3. $P(S \text{ on trial } i) = p$ for all $i = 1, 2, 3, \dots$
fixed
4. The experiment continues until a total of r successes have been observed, where r is a specified positive integer.

The random variable of interest is X , the number of failures that precede the r th success; X is called a **negative binomial random variable**.

PROPOSITION The pmf of the negative binomial rv X with parameters $r = \text{number of } S\text{'s}$ and $p = P(S)$ is

$$nb(x; r, p) = \binom{x+r-1}{x} p^r \cdot (1-p)^x \quad \binom{x+r-1}{x} = \binom{x+r-1}{r-1}$$

x F's $r-1$ S's

and mean and variance are

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

EXAMPLE A couple want to have 2 girls. Find the probability that they have 5 children to get 2 girls.

g, b, b, b

— — — — g

Let ~~X~~ track # of boys among 5 children.

$$nb(3:2, 0.5) = \binom{4}{3} \cdot \left(\frac{1}{2}\right)^2 \left(1 - \frac{1}{2}\right)^3 = 4 \cdot \left(\frac{1}{2}\right)^5$$

EXAMPLE A family decides to have children until it has three children of the same gender. Assuming $P(B) = P(G) = 0.5$, what is the pmf of X = the number of children in the family?

$$\text{range}(X) = \{3, 4, 5\}$$

$$P(X=3) = \frac{1}{4}$$

$$\underline{b} \underline{b} \underline{b} \text{ or } \underline{g} \underline{g} \underline{g} \quad 2 \times \left(\frac{1}{2}\right)^3 = \frac{1}{4}$$

$$\begin{array}{c} 2 \text{ b's, } g \\ \underline{\quad} \underline{\quad} \underline{\quad} \underline{b} \end{array}, \quad \underline{\quad} \underline{\quad} \underline{\quad} \underline{g}$$

$$3 \text{ b's} \quad 2 \binom{3}{1} \cdot \left(\frac{1}{2}\right)^4 = 3 \text{ g's} \quad \frac{3}{8}$$

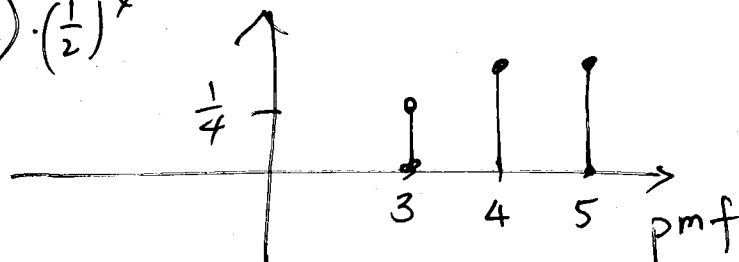
2 b's 2 g's

$$\underbrace{\underline{\quad} \underline{\quad} \underline{\quad} \underline{\quad}}_{3 \text{ b's}} \underline{b} \quad 2 \times \binom{4}{2} \cdot \left(\frac{1}{2}\right)^5 = 6 \cdot \frac{1}{16} = \frac{3}{8}$$

3 b's

$$\binom{4}{2} \cdot \left(\frac{1}{2}\right)^4$$

X	$P(X)$
3	$\frac{1}{4}$
4	$\frac{3}{8}$
5	$\frac{3}{8}$
6	0
7	0
8	0



Practice exercises in Section 3.5: all odds from 69 to 77.

3.6 The Poisson probability distribution

In contrast to binomial, hypergeometric, and negative binomial distributions where they are all derived by starting with an experiment consisting of trials or draws from which we apply the laws of probability to find their distributions, the **Poisson** distribution~~x~~ is not based on any simple experiment, but instead by certain limiting operations.

DEFINITION A discrete random variable X is said to have a **Poisson distribution**~~x~~ with parameter μ , ($\mu > 0$) if the pmf of X is

$$p(x; \mu) = \frac{e^{-\mu} \mu^x}{x!},$$

poisson pdf (μ, x)

for $x = 0, 1, 2, 3, \dots$

Does $p(x; \mu)$ specify a legitimate pmf? Let us check.

1. positivity

$$p(x; \mu) \geq 0 \quad x \in \{0, 1, 2, \dots\}.$$
$$e^{-\mu} \cdot \frac{\mu^x}{x!}$$

2. sum is 1. Use Taylor series for e^x expanded about $x = 0$.

$$\sum_{n=0}^{\infty} e^{-\mu} \frac{\mu^n}{n!} = e^{-\mu} \underbrace{\sum_{n=0}^{\infty} \frac{\mu^n}{n!}}_{e^{\mu}} = e^{-\mu} \cdot e^{\mu} = 1.$$

Some typical situations are:

1. Probability an earth quake of magnitude greater than 2.5 will hit Vancouver tomorrow.
2. Probability that a big earth quake will hit Vancouver this year.

3. Probability that at least one fatal car accident will occur on the Sea-to-Sky highway this month.
4. Probability that a skier who goes out of bounds will need to be rescued this winter.

Events that occur *randomly* in time like earth quakes, accidents, computer system crashes, or in space like the number of flaws in a roll of fabric, the number of weak spots in a length of piping where a known average rate is available are often modelled by the Poisson distribution.

PROPOSITION Suppose that in the binomial pmf $b(x; n, p)$, we let $n \rightarrow \infty$ and $p \rightarrow 0$ in such a way that np approaches a value $\mu > 0$. Then

$$\underline{b(x; n, p)} \longrightarrow p(x; \mu) \quad \mu = np$$

Conservative OK to approximate a binomial pmf by the Poisson pmf whenever $n \geq 100$, $p \leq 0.01$ and $np \leq 20$. This is more relaxed than the guidelines given in the text of $np < 5$ & $n > 50$.

The mean and variance of a rv X with a Poisson distribution with parameter μ is also μ .

EXAMPLE Suppose that only 0.10% of all computers of a certain type experience CPU failure during the warranty period. Consider a sample of 10 000 computers.

1. What are the expected value and standard deviation of the number of computers in the sample that have the defect? *cpu*

Let success be the event computer fails.

$$P(S) = 0.1\% \quad , \quad n = 10,000$$

$$b(x; n, p) \approx E(X) = 10000 \times \frac{1}{1000} = 10 \quad \sigma_X = \sqrt{V(X)} = \sqrt{np(1-p)}$$

2. What is the (approximate) probability that more than 10 sampled computers have the defect?

$$\approx 3.16$$

OK to use Poisson, $\mu = 10$.

$$P(X \geq 10) = 1 - \underbrace{P(X \leq 10)}_{\text{cdf}} \approx 1 - F(10; 10) \approx 0.4170$$

approx with poisson cdf
poisson cdf(μ, x)

3. What is the (approximate) probability that no sampled computer has the defect?

$$P(X=0) = p(0; 10) = \text{poisson pdf}(10, 0) \approx 0.00004540 \approx 4.540 \times 10^{-5}$$

The mean and variance of a rv X with a Poisson distribution with parameter μ is also μ .

The Poisson process

Application of the Poisson distribution: occurrence of events of some type over time like visits to a particular website, pulses of some sort recorded by a counter, email messages sent to a particular address, accidents in an industrial facility, etc. The following assumptions about the way in which the events of interest occur accompany the Poisson distribution.

1. There exists a parameter $\alpha > 0$ such that for any short time interval of length Δt , the probability that exactly one event occurs is

$$\alpha \cdot \Delta t + o(\Delta t) \quad \lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0$$

2. The probability of more than one event occurring during Δt is

$$o(\Delta t), \text{ so prob of no event during } \Delta t \text{ is } 1 - \alpha \cdot \Delta t - o(\Delta t)$$

3. The number of events occurring during the time interval Δt is independent of the number that occur prior to this time interval.

"memory less process"

PROPOSITION $P_k(t) = e^{-\alpha t} \cdot \frac{(\alpha t)^k}{k!} \leftarrow \begin{matrix} \text{\# of events (k)} \\ \mu = \alpha t \end{matrix}$

during a time interval of length t

$E(\# \text{ of events})$ during such time interval is αt , so expected # during a unit interval of time is α , called the rate of process.

Other common variables usually modelled by the Poisson process are

1. the number of customers arriving at a service facility,
2. the number of telephone calls per hour at a switch board,
3. the number of hits on a website,
4. the number of cars passing through an intersection.

EXAMPLE The number of people arriving for treatment at Burnaby General Hospital's emergency room can be modelled by a Poisson process with rate parameter of five (per hour). $\alpha = 5/\text{hr}$

1. What is the probability that exactly four arrivals occur during a particular hour?
 $t = 1 \text{ hr}$

$$P_4(1) = e^{-5 \cdot 1} \cdot \frac{(5 \cdot 1)^4}{4!} = \text{poisson pdf}(5, 4) \approx 0.1755$$

μ, x
 $K=4$
2. What is the probability that at least four people arrive during a particular hour?

$$P(X \geq 4) = 1 - P(X \leq 3)$$

$$= 1 - \text{poisson cdf}(5, 3)$$

$$= 1 - F(3; 5)$$

$$\approx 0.7350$$

86

3. How many people do you expect to arrive during a 45-minute period?

$$t = \frac{3}{4} \text{ hr.}$$

$$E(X) = \frac{15}{4}$$

$$\mu = \lambda t = 5 \times \frac{3}{4}$$

5 Legal sized pages.

chap 1 $\frac{1}{5}$

chap 2. . count

Prob

Bayes

reliability

23

Chap 3: graphs pmf vs cdf. $\frac{1}{6}$

distributions

$\frac{1}{8}$

Applications

$\frac{1}{14}$

$\frac{1}{56}$



Practice exercises in Section 3.6: 79, 81, 83, 85, 91.

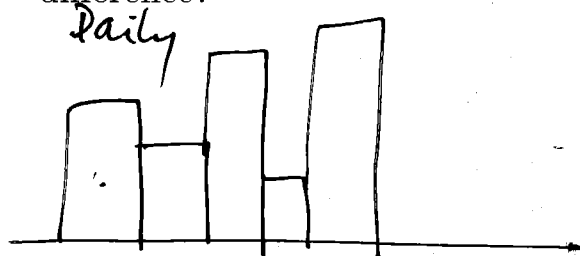
Chapter 4: Continuous Random Variables & Probability Distributions

4.1 Probability density functions

Recall discrete vs. continuous random variables:

- A discrete random variable X
possible values are finite or
can be listed in an infinite sequence.
- A continuous random variable X
values from an interval or
union of disjoint intervals.

EXAMPLE Rainfall in Greater Vancouver during February. Try to record daily, every 12 hours, every 3 hours, every hour. What is the difference?

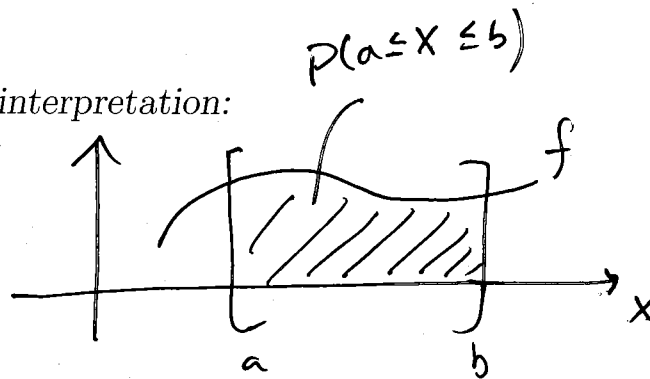


$t \rightarrow 0$
Riemann sum.

DEFINITION Let X be a continuous rv. Then a **probability distribution** or **probability density function** (pdf) of X is a function $f(x)$ such that for any two numbers a and b with $a \leq b$,

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

Graphical interpretation:



Conditions for f to be a legitimate pdf:

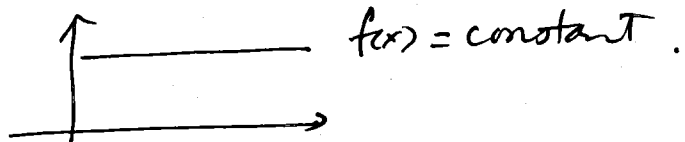
1. non-negativity

$$f(x) \geq 0 \quad \text{for all } x \in \text{domain of } f.$$

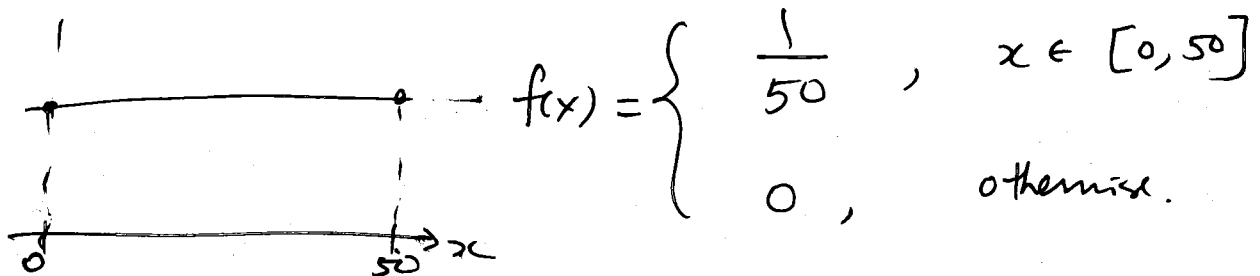
2. totality

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

Uniform distribution



EXAMPLE From our classroom to the bus station is 50 metres. Let X be the distance from our classroom on the path to the bus station. Your falling at position X is subject to uncertainty. One possible pdf for X is:



1. Find the probability that you fall before midway.
2. Find the probability that you fall between 20 m and 40 m from our classroom.