MATH 201 (SUMMER 2023, SESH A2)

LECTURE 6 : 05/23/23

ANURAG SAHAY
OFF HRS: BY APPT (VIA ZOOM)

email: anuragsahay@rochester.edu

Zoom ID: 979-4693-0650

LECTURES:
9:00 AM – 11:15 AM (ET)
M, T, W, R

COURSE WEB PAGE: https://people.math.rochester.edu/grads/asahay/summer2023/math201/index.html

ALL PHOTOS TAKEN FROM TEXTBOOK

ANNOUNCEMENTS

1. LECTURE 5, WEEK 2 H.W. IS UPLOADED. (PANOPTO/WEBSITE)
2. OFFICE HOURS: TR, 11:15 AM - 12:15 PM ET. (TUES/THURS RIGHT AFTER CLASS)
3. DEADLINES: AT 11 PM.
   - a. WW03 - TUES, MAY 23rd
   - b. HW02 - TUES, MAY 23rd
   - c. WW04 - FRI, MAY 26th
   - d. HW03 - SAT, MAY 27th

   TO BE UPLOADED
4. NO IN-CLASS LECTURE ON WEDNESDAY (MAY 24th) → WILL BE RECORDED.
5. PLEASE KEEP YOUR VIDEOS ON, IF POSSIBLE!

§ 3.1 PROBABILITY DISTRIBUTIONS OF RVs

TO DESCRIBE X (R.V.) COMPLETELY,

NEED: $P(X \in B)$ $B \subseteq \mathbb{R}$ → DISTRIBUTION FUNCTION OF X.

LESS INFO CAN BE ENOUGH:

① p.m.f. (cf § 1.5) → PROBABILITY MASS F (DISCRETE)

② p.d.f. → PROBABILITY DENSITY F (CONT.)

(cf § 3.2)

③ c.d.f. → CUMULATIVE DISTRIBUTION FUNCTION

RECALL $\quad p.m.f \equiv p(\cdot) = p_X(\cdot)$

IF $\quad X \in \underbrace{A}_{\text{DISCRETE}} = \{ k_1, k_2, \ldots \}$

$$p_X(k) = p(k) = P(X = k)$$

THEN,

$$P(X \in B) = \sum_{k \in B} p_X(k)$$

FINITE OR COUNTABLE SUM.

$$\& \quad \sum_{k} p_X(k) = \sum_{k \in A} p_X(k) = 1$$

p.d.f

WHEN p.d.f. EXISTS WE SAY
X IS CONTINUOUS.

Definition 3.1. Let X be a random variable. If a function f satisfies

$$P(X \le b) = \int_{-\infty}^{b} f(x)\, dx \tag{3.3}$$

for all real values b, then f is the **probability density function** (p.d.f.) of X.

$\rightarrow f_X(x)$

$$P(X \le b) = P(X \in \underbrace{(-\infty, b]}_{(a,b]\ ?})$$

$(0, \infty)$?

FACTS: ① $P(X \in B) = \int_B f(x)\, dx$ FOR "REASONABLE" $B \subseteq \mathbb{R}$

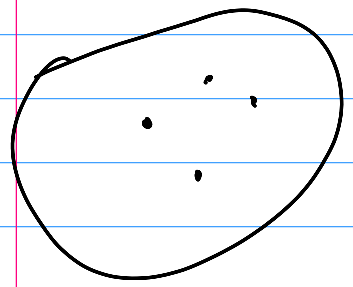
e.g. $P(X \in (a,b)) = \int_a^b f(x)\, dx$ $\rightarrow$ COMPLETELY DET. DISTRIBUTION.

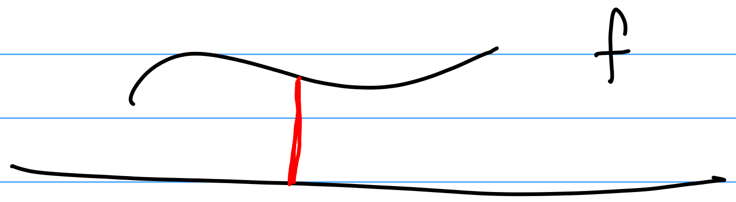
② NO POINT MASSES :

$X \to$ cont. R.V.

$$P(X=c) = P\big(X \in [c,c]\big) = \int_{[c,c]} f(x)\,dx = \int_c^c f(x)\,dx = 0$$

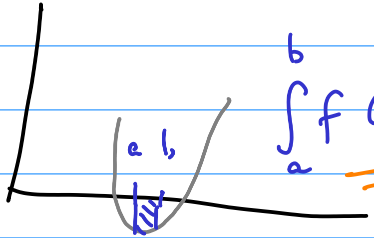
[$c \in \mathbb{R}$, FIXED]





NOT DISCRETE !

③ $f(x) \geq 0 \quad \forall x \in \mathbb{R}$



$\int_a^b f(x)\,dx < 0 \rightarrow \#$

$\hookrightarrow P(a \leq X \leq b) \geq 0$

$\rightarrow *$

④ $\int_{-\infty}^{\infty} f(x)\,dx = P(X \in \mathbb{R}) = 1$

$\hookrightarrow$ COMPARE TO p.m.f.

$\sum_k p_X(k) = P(X \in \mathbb{R}) = 1$

ANY FUNCTION SATISFYING $*$ IS A p.d.f. OF SOME X.

$f(x) = \frac{1}{\text{LENGTH } [a,b]}$ $x \in [a,b]$

0 o.w.

$P(I) \propto \text{LENGTH}(I)$

$P[a,b] = 1$

$P[c,d] = \frac{d-c}{b-a}$

$1/b-a$

$\mathbb{R}$

a b

Definition 3.5. Let $[a, b]$ be a bounded interval on the real line. A random variable X has the *uniform distribution on the interval* $[a, b]$ if X has density function

$$f(x) = \begin{cases} \boxed{\frac{1}{b-a},} & \text{if } x \in [a,b] \\ 0, & \text{if } x \notin [a,b]. \end{cases} \tag{3.8}$$

Abbreviate this by $X \sim \text{Unif}[a, b]$.

Q. WHAT IS $P(X \in [c,d]) = \int_c^d f(x)\,dx = \int_c^d \left(\frac{1}{b-a}\right) dx$

$a \le c \le d \le b$

$$= \frac{x}{b-a}\Big]_c^d = \frac{d-c}{b-a}$$

NOTE: POINTS HAVE NO MASS

$$\Rightarrow \quad P(c \le X \le d) = \underbrace{P(X = c)}^{0} + P(c < X \le d)$$

$$\left[c \le X \le d = (X = c) \cup (c < X \le d) \right]$$

$$P(c \le X \le d) = P(c < X \le d)$$
$$= P(c \le X < d)$$
$$= P(c < X < d)$$

$[a,b]$, $(a,b]$, $[a,b)$, (a,b) $\} \rightarrow$ EQUAL PROBABILITY.

Example 3.6. Let Y be a uniform random variable on $[-2, 5]$. Find the probability that its absolute value is at least 1.

$Y \sim \text{Unif}\,[-2, 5]$

$$f(x) = \begin{cases} \dfrac{1}{5-(-2)} = \dfrac{1}{7} & x \in [-2, 5] \\ 0 & x \notin [-2, 5] \end{cases}$$

$P(\underbrace{|Y| \geq 1}_{Y \in B})$

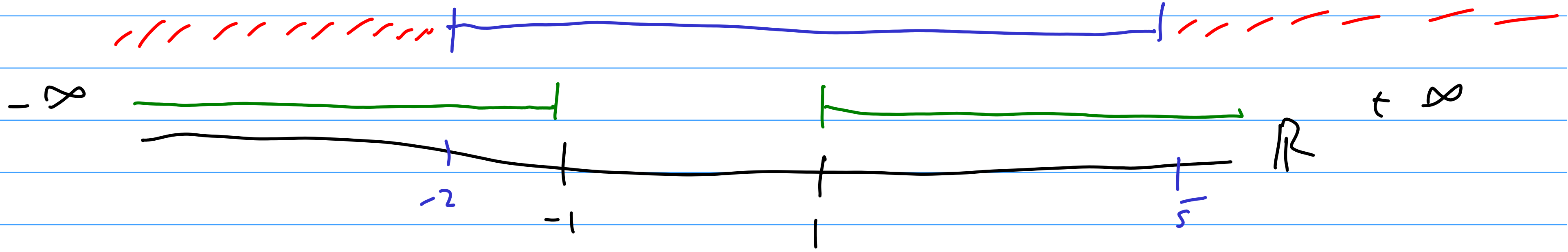
$|Y| \geq 1 \iff Y \geq 1$ or $Y \leq -1$

$\iff Y \in (-\infty, -1] \cup [1, \infty)$

$$P(|Y| \geq 1) = P(Y \in \underline{(-\infty, -1]} \cup \underline{[1, \infty)})$$

$$= P(Y \in (-\infty, -1]) + P(Y \in [1, \infty))$$

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



$$= \int_{-2}^{-1} f(x)\,dx + \int_{1}^{5} f(x)\,dx = \frac{1}{7} + \frac{4}{7} = \frac{5}{7}$$

N.B. : $f(x)$ IS NOT A PROBABILITY.

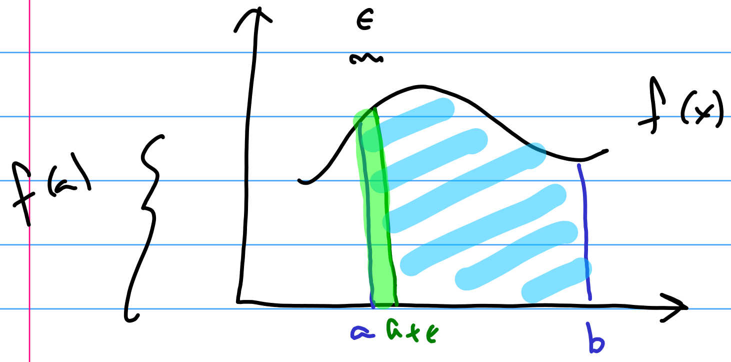
(UNLIKE p.m.f.)

HOWEVER, $f(x) \cdot \epsilon \simeq$ A PROB.

CALCULUS.

Fact 3.7. Suppose that random variable X has density function f that is continuous at the point a. Then for small $\varepsilon > 0$

$$P(a < X < a + \varepsilon) \approx f(a) \cdot \varepsilon.$$



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

$P(a < X < a+\epsilon) \simeq f(x) \cdot \epsilon$

$$\lim_{\epsilon \to 0} \frac{P(a < X < a+\epsilon)}{\epsilon} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_a^{a+\epsilon} f(x)\,dx = \frac{d}{dt} \int_a^{t} f(x)\,dx \Bigg|_{t=a}$$

$$= f(a)$$

|||ly

$$P(\overbrace{a-\epsilon < X < a}) \approx f(x) \cdot \epsilon$$

$$P(\underbrace{a-\epsilon < X < a+\epsilon}) \approx f(x) \cdot 2\epsilon$$

Example 3.8. Suppose the density function of X is $f(x) = 3x^2$ for $0 < x < 1$ and $f(x) = 0$ elsewhere. Compare the precise value of $P(0.50 < X < 0.51)$ with its approximation by Fact 3.7.

NOTE : $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 3x^2\,dx$

$= x^3\Big]_0^1 = 1$

$$P(0.5 < X < 0.51) = \int_{0.5}^{0.51} 3x^2\,dx$$

$$= x^3\Big]_{0.5}^{0.51} = (0.51)^3 - (0.5)^3$$

$$= 0.007651$$

$\epsilon = 0.01$ $a = 0.5$

$P(0.5 < X < 0.51) \approx f(0.5) \times 0.01$

$(f(a)\,\epsilon)$

$$= 3 \times (0.5)^2 \times 0.01 = 0.0075$$

§ 3.2 CUMULATIVE DISTRIBUTION FUNCTION

c.d.f.

$P \leftarrow$ PROB.

$p_X \leftarrow$ p.m.f. — DISCRETE

$f_X \leftarrow$ p.d.f. — CONTINUOUS

$F_X \leftarrow$ c.d.f. — ANY (INCL. NON-DISCRETE & CONT.)

Definition 3.10. The cumulative distribution function (c.d.f.) of a random variable X is defined by

$$F(s) = P(X \leq s) \quad \text{for all } s \in \mathbb{R}. \tag{3.12}$$

$$P(\underset{\text{OPEN}}{a <} X \underset{\text{CLOSED}}{\leq} b) = P\left(\begin{array}{c} X \leq b \\ \text{BUT} \quad \text{NOT} \quad X \leq a \end{array}\right) = \overbrace{P(X \leq b)}^{F(b)} - \overbrace{P(X \leq a)}^{F(a)}$$

$(\because X \leq a \Rightarrow X \leq b)$

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

FACT: KNOWING c.d.f. IS ENOUGH TO DETERMINE DISTRIBUTION $P(X \in B)$

c.d.f. OF A DISCRETE RV

$$F(s) = \underbrace{P(X \leq s)}_{X \in (-\infty, s]} = \sum_{k \leq s} \underbrace{P(X = k)}_{\text{p.m.f.}}$$

Example 3.11. Let $X \sim \text{Bin}(\underset{n}{2}, \underset{=p}{\tfrac{2}{3}})$ (recall Definition 2.32). Find the cumulative distribution function of X. AND PLOT

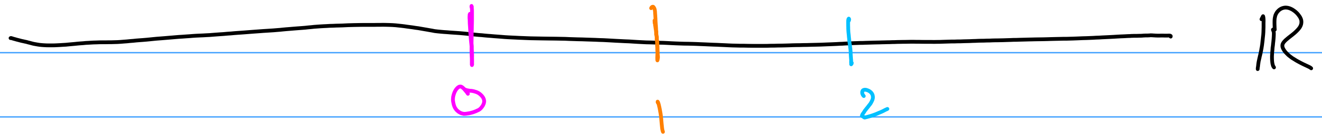
$\{0, 1, \ldots, n\}$, $n = 2$

$$P(X = k) = \binom{2}{k} \left(\frac{2}{3}\right)^{k} \left(1 - \frac{2}{3}\right)^{2-k}$$

$$\binom{n}{k} p^{k} (1 - p)^{n-k}$$

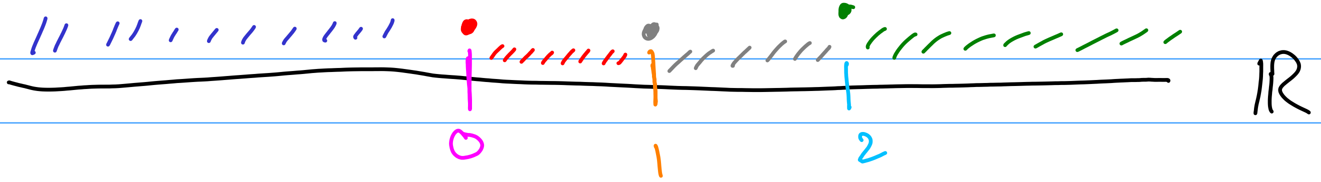
$$p_X(k) = P(X=k) = \begin{cases} 1/9 & k = 0 \\ 4/9 & k = 1 \\ 4/9 & k = 2 \\ 0 & \text{o.w.} \end{cases}$$

$$F_X(s) = P(X \leq s) = \sum_{k \leq s} P(X = k)$$

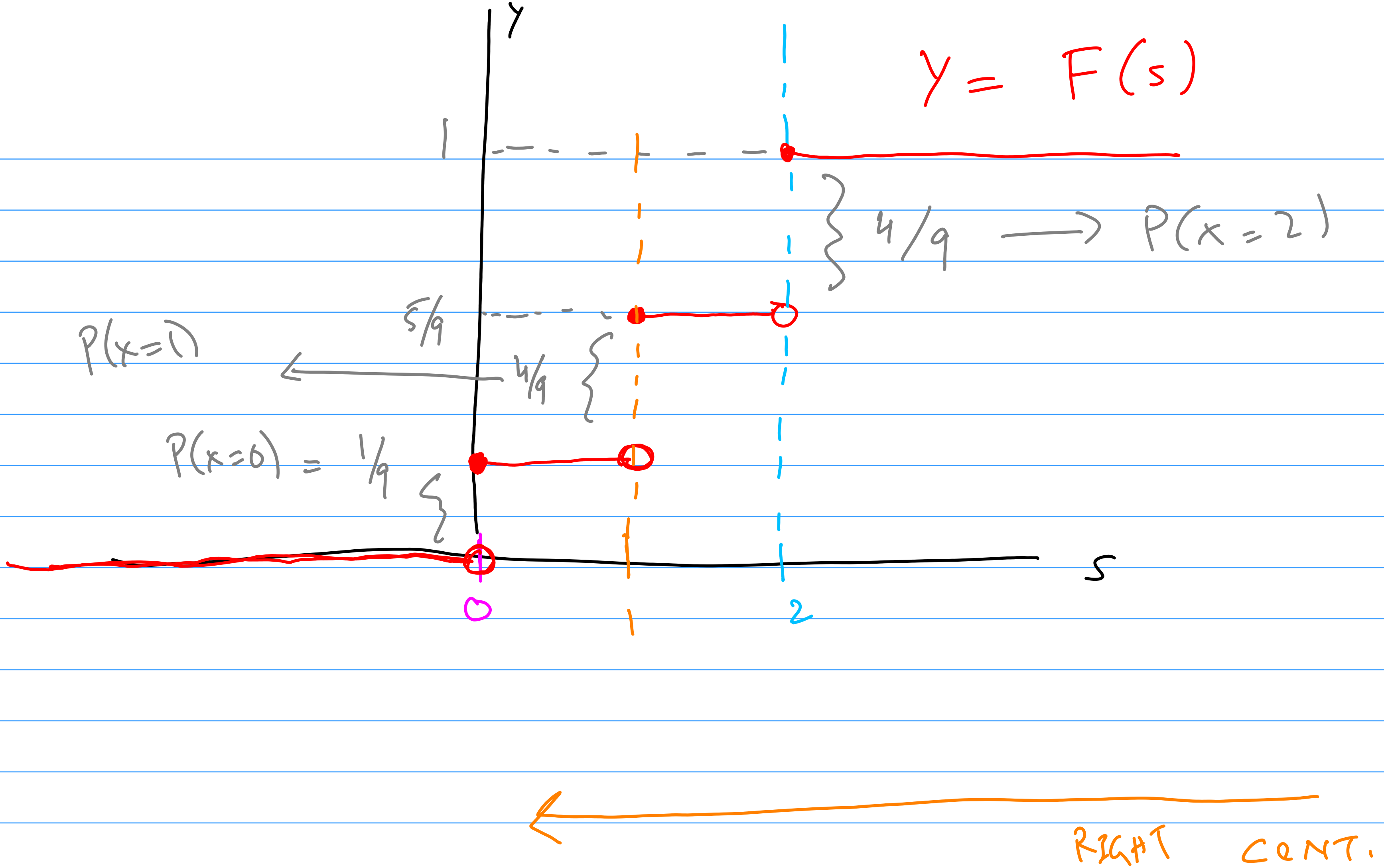


$(k \in \{0, 1, 2\})$

$$F_X(s) = P(X \le s) = \sum_{k \le s} P(X = k)$$



$$F_X(s) = \begin{cases} 0 & \text{IF } s < 0 \\ P(X=0) = 1/9 & \text{IF } 0 \le s < 1 \\ P(X=0) + P(X=1) = 5/9 & \text{IF } 1 \le s < 2 \\ P(X=0) + P(X=1) + P(X=2) = 9/9 = 1 & \text{IF } s \geq 2 \end{cases}$$



y
y = F(s)
1
4/9
P(x=2)
5/9
P(x=1)
4/9
P(x=0) = 1/9
s
0
1
2
RIGHT CONT.

POINT OF DISCONTINUITY

NOTES: (1) IF X IS DISCRETE. F_X IS A STEP FUNCTION. → $\boxed{\text{JUMPS}}$ AT VALUES X TAKES.

(2) $P(X = k)$ = SIZE OF JUMP OF F_X AT k

(3) $F_X(s) = P(X \underbrace{\leq}_{\text{CRUCIAL.}} s)$

(e.g. $F(0) = 1/9 \leftrightarrow P(X < 0) = 0$

BREAK TILL

10:10 AM

c.d.f. OF A CONTINUOUS R.V.

$$F(s) \underbrace{=}_{\text{defn}} P\underbrace{(X \le s)}_{x \in (-\infty, s]} \overbrace{=}^{\text{CONT.}} \int_{-\infty}^{s} f(x)\,dx$$

$\Rightarrow F'(s) = f(s)$ (F.T.C. ; IF f CONT. AT s)

↳ FUND. THM. OF CALCULUS)

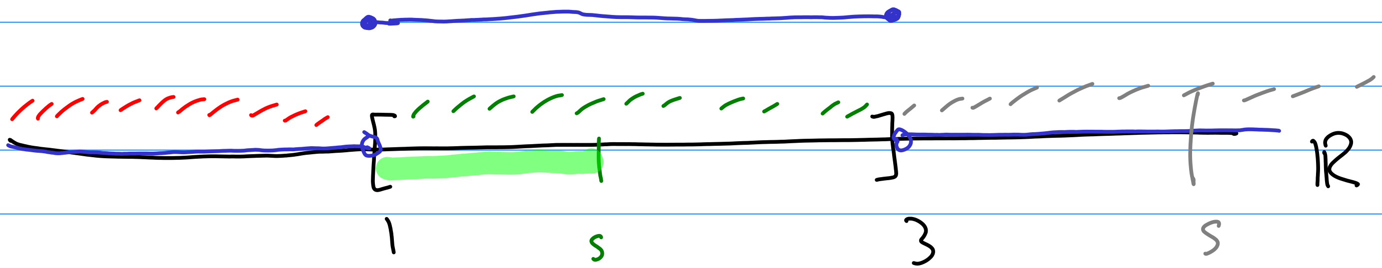
Example 3.12. Let X be a Unif[1, 3] random variable. Find the cumulative distribution function of X. AND PLOT

$$f(x) = \begin{cases} \frac{1}{3-1} = \frac{1}{2} & x \in [1,3] \\ 0 & x \notin [1,3] \end{cases}$$

$$\Downarrow$$

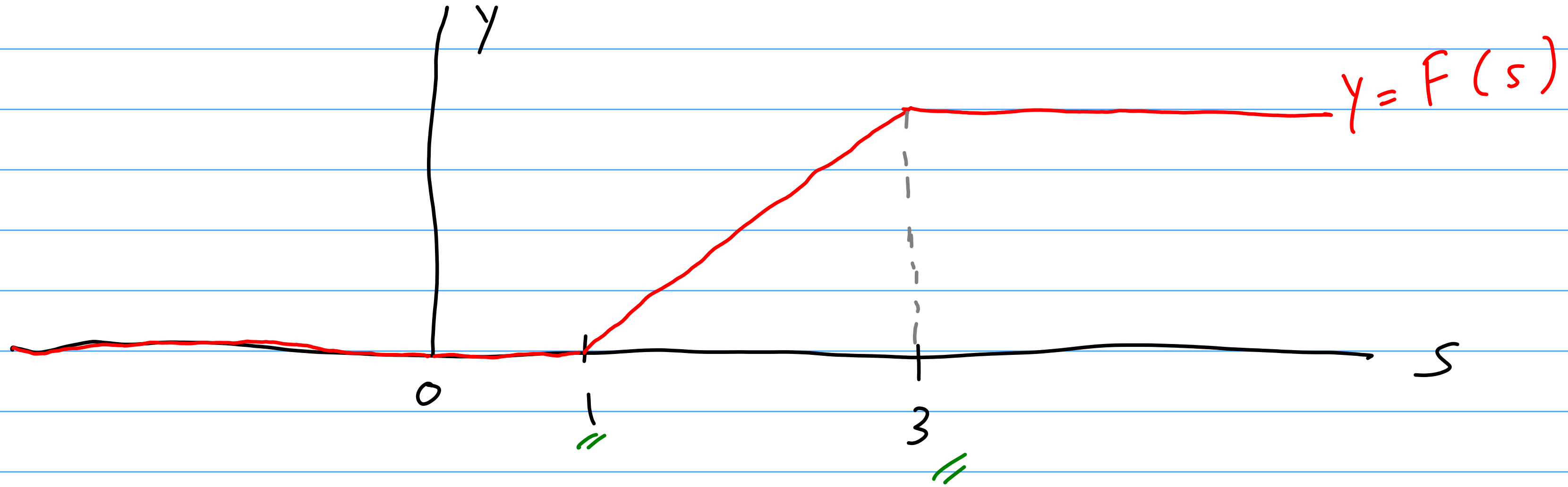
$$f(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{2} & 1 \le x \le 3 \\ 0 & x > 3 \end{cases}$$

p.d.f.



$$F(s) = \int_{-\infty}^{s} f(x)\,dx = \begin{cases} 0 & (s<1) \\ \int_1^s \left(\tfrac{1}{2}\right) dx = \dfrac{s-1}{2} & (1 \le s \le 3) \\ \int_1^3 \left(\tfrac{1}{2}\right) dx = 1 & (s>3) \end{cases}$$

$$F_X(s) = \begin{cases} 0 & s < 1 \\ \frac{s-1}{2} & 1 \le s \le 3 \\ 1 & s > 3 \end{cases}$$



NOTE:

1. CONTINUOUS (BECAUSE X IS CONT.)

2. $F'(s) = f(s)$ (ALMOST EVERYWHERE)

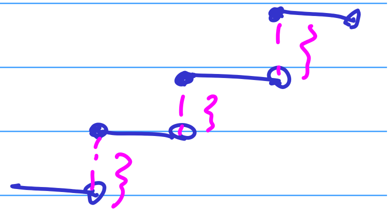
FINDING p.m.f. / p.d.f. FROM c.d.f

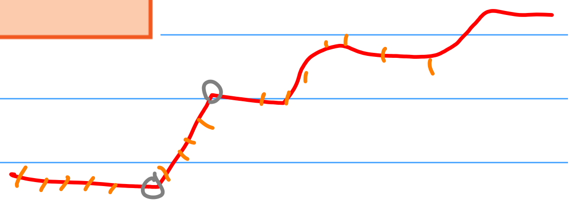
STEP

Fact 3.13. Let the random variable X have cumulative distribution function F.

(a) Suppose F is piecewise constant. Then X is a discrete random variable. The possible values of X are the locations where F has jumps, and if x is such a point, then $P(X = x)$ equals the magnitude of the jump of F at x.

(b) Suppose F is continuous and the derivative $F'(x)$ exists everywhere on the real line, except possibly at finitely many points. Then X is a continuous random variable and $\boxed{f(x) = F'(x)}$ is the density function of X. If F is not differentiable at x, then the value $f(x)$ can be set arbitrarily. ♣





PROP. OF c.d.f.

Fact 3.16. Every cumulative distribution function F has the following properties.

(i) Monotonicity: if $s < t$ then $F(s) \le F(t)$. $\rightarrow P(X \le t)$ $P(X \le s)$

[FOLLOWS PROB. FROM ARE MONOTONE]

(ii) Right continuity: for each $t \in \mathbb{R}$, $F(t) = \lim_{s \to t^+} F(s)$ where $s \to t^+$ means that s approaches t from the right.

NO PROOF.

(iii) $\lim_{t \to -\infty} F(t) = 0$ and $\lim_{t \to \infty} F(t) = 1$.

$$\lim_{t \to -\infty} F(t) = P(X \le -\infty) = P(\emptyset) = 0$$

$$X \le s \subseteq X \le t$$

$$\lim_{t \to \infty} F(t) = P(-\infty < X < \infty) = P(\Omega) = 1$$

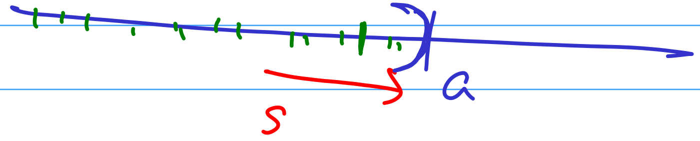
$$F(a^-) = \lim_{s \to a^-} F(s)$$

Fact 3.17. Let X be a random variable with cumulative distribution function F. Then for any $a \in \mathbb{R}$

$$P(X < a) = \lim_{s \to a^-} F(s). \quad \neq F(a) \qquad (3.16)$$

(IN GENERAL)

$$F(s) = P(X \le s) \xrightarrow[s \to a^-]{} P(X < a)$$



CORR:

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

$$= F(a) - F(a^-) \longrightarrow 0 \text{ WHEN } F \text{ CONT. AT } a.$$

$$= \text{JUMP AT } a$$

MASS AT a (NOT NECESSARILY DISCRETE)

DISCONTINUITIES OF F $\longleftrightarrow$ POINT MASSES OF THE DISTRIBUTION

IN PARTICULAR

F IS CONT. $\leftrightarrow$ NO POINT MASSES.

F' EXISTS A.E. $\leftrightarrow$ p.d.f.

Example 3.20. Carter has an insurance policy on his car with a \$500 deductible. This means that if he gets into an accident he will personally pay for 100% of the repairs up to \$500, with the insurance company paying the rest. For example, if the repairs cost \$215, then Carter pays the whole amount. However, if the repairs cost \$832, then Carter pays \$500 and the remaining \$332 is covered by the insurance company.

Suppose that the cost of repairs for the next accident is uniformly distributed between \$100 and \$1500. Let X denote the amount Carter will have to pay. Find the cumulative distribution function of X. AND PLOT

~~e.g.~~: COST OF REPAIRS = \$800

CARTER $\rightarrow$ \$500, INSURANCE \$800 − \$500 = \$300.

Y = COST OF REPAIRS. $Y \sim \text{Unif}[100, 1500]$

X = AMOUNT CARTER PAYS.

$$Y \sim \text{Unif}\,[100,\ 1500]$$

p.d.f.

$$f_Y(t) = \begin{cases} 0 & t < 100 \\ \dfrac{1}{1500-100} = \dfrac{1}{1400} & t \in [100, 1500] \\ 0 & t > 1500 \end{cases}$$

c.d.f.

$$F_Y(s) = \begin{cases} 0 & s < 100 \\ (s-100)/1400 & 100 \le s \le 1500 \\ 1 & s > 1500 \end{cases}$$

$s < 100$

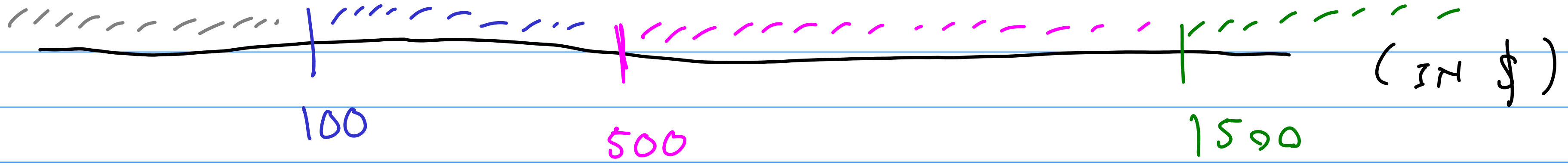
$$f_X(s) = P(X \le s)$$
$$\le P(X < 100)$$
$$= 0$$

$$F_X(s) = P(X \le s)$$
$$= P(Y \le s)$$
$$= \frac{s-100}{1500-100}$$
$$= \frac{s-100}{1400}$$

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

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

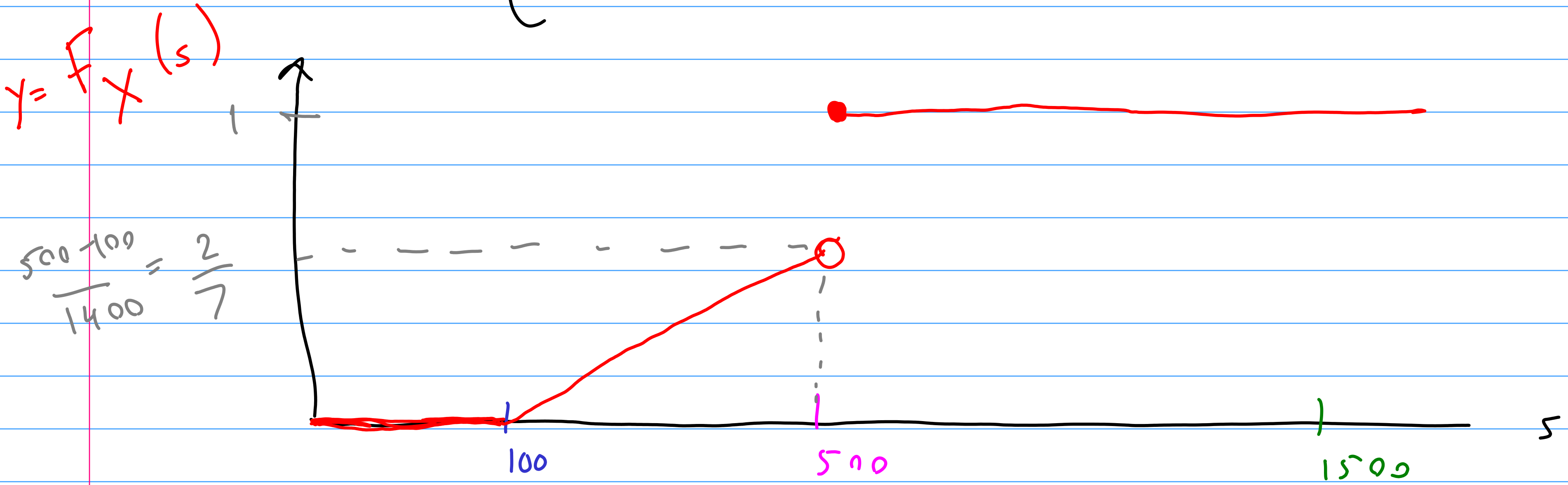
(CARTER'S LIABILITY IS ALWAYS ≤ 500)

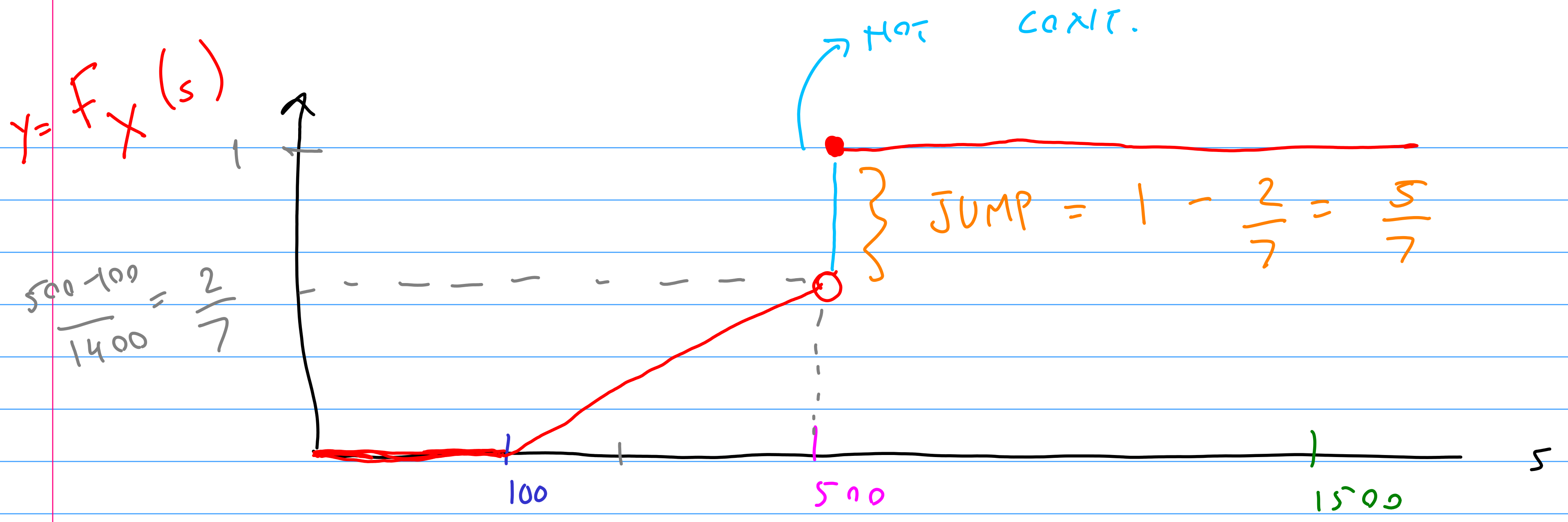


$X \to$ CARTER'S LIABILITY

$Y \to$ ACTUAL COST

$$F_X(s) = \begin{cases} 0 & s < 100 \\ F_Y(s) = \dfrac{s-100}{1400} & s \in [100, 500) \\ 1 & s \geq 500 \end{cases}$$





(1) DISCRETE? → NO.

(2) CONT. ? → NO.

Fact 3.13. Let the random variable X have cumulative distribution function F.

(a) Suppose F is $\boxed{\text{piecewise constant}}$ → STEP. Then X is a discrete random variable. The possible values of X are the locations where F has jumps, and if x is such a point, then $P(X = x)$ equals the magnitude of the jump of F at x.

(b) Suppose $\boxed{F \text{ is continuous}}$ and the $\boxed{\text{derivative } F'(x)}$ exists everywhere on the real line, except possibly at finitely many points. Then X is a continuous random variable and $f(x) = F'(x)$ is the density function of X. If F is not differentiable at x, then the value $f(x)$ can be set arbitrarily. ♣

$P(X = 500) =$ VALUE OF JUMP AT 500

$= \frac{5}{7}$

$\frac{2}{7} \to P(\underbrace{100 \le X < 500}_{\text{NO POINT MASS.}})$

$$P(X=500) = P(\underset{\substack{\downarrow \\ \text{TOTAL COST}}}{Y} \geq 500) = \int_{500}^{\infty} f_Y(t)\,dt$$

(RECALL $Y \sim Unif[100, 1500]$) $\frac{1}{1500-100}$

$$\Rightarrow P(X=500) = \int_{500}^{1500} \frac{1}{1400}\,dt$$

$$= \frac{1500-500}{1400} = \frac{1000}{1400} = \frac{5}{7} \;//$$

REMINDER: IN-PERSON

1. NO CLASS TOMORROW

   RECORDING LATER TODAY

2. OFFICE HOURS