$\boxed{\text{Jan 19}}$ Quiz. Solutions to quizzes, Midterm Exams posted; HW & Final not.

Remarks 1. If f is a density then

- $f \geq 0$ pointwise (since $f = F' \geq 0$ as F is $\uparrow$).
- $\int_{-\infty}^{\infty} f(x)\,dx = 1$ $\left(= \lim_{a\to\infty} F(a) = 1 \right)$.
- $\mathbb{P}\{a < X < b\} = \int_a^b f(x)\,dx$.

2. If X is abs. continuous then

$$\mathbb{P}\{X = x\} = 0 \qquad \forall x \in \mathbb{R}.$$

$$\left(= F(x) - F(x-) = 0 \text{ by continuity of } F \right).$$

3. $\exists$ r.v.'s that are neither discrete nor continuous. (Why?)

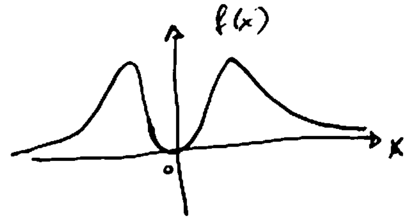
Functions of r.v.'s: Change of density

~~Ex (Change of density)~~

Ex. 1. Let $X \sim N(0,1)$; $\quad Y := X^3$. $\quad$ density of $Y = ?$

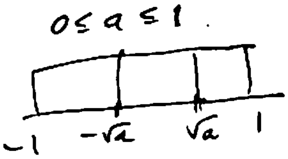
CDF: $F_Y(a) = \mathbb{P}\{Y \leq a\} = \mathbb{P}\{X^3 \leq a\} = \mathbb{P}\{X \leq a^{1/3}\} = F_X(a^{1/3})$.

Density: $f_Y(a) := \frac{d}{da} F_Y(a) = f_X(a^{1/3}) \cdot \frac{1}{3}|a|^{-2/3} = \frac{1}{3|a|^{2/3}\sqrt{2\pi}}\, e^{-\frac{|a|^{2/3}}{2}}, \quad a \neq 0$

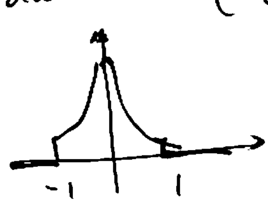


Ex. 2 Let $X \sim \text{Unif}[-1,1]$, $\quad Y = X^2$.

CDF: $F_Y(a) = \mathbb{P}\{Y \leq a\} = \mathbb{P}\{X^2 \leq a\} = \mathbb{P}\{|X| \leq \sqrt{a}\} = \sqrt{a}, \quad 0 \leq a \leq 1$.



$$f_Y(a) = \frac{d}{da} F_Y(a) = \begin{cases} \frac{1}{2\sqrt{a}}, & 0 < a \leq 1 \\ 0, & \text{—} \end{cases}$$



2.5. Expectations of discrete r.v's.

Def. Let X be a discrete r.v. with PMF $p(x_k) = P\{X = x_k\}$.
The <u>expectation</u> (expected value) of X is defined as

$$\mathbb{E}[X] = \sum_k x_k \, p(x_k).$$

We require that this series converges <u>absolutely</u>;
otherwise we say that X has no expectation.

Ex. $X = \#$ heads in 3 coin flips:

$$\mathbb{E}[X] = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = 1.5$$

Remark If all N values of X are equally likely then $p(x_k) = \frac{1}{N} \Rightarrow$

$$\mathbb{E}[X] = \frac{1}{N} \sum_1^N x_k \quad \text{(arithmetic mean)}$$

Otherwise, weighted by $p(x_k)$.

Ex (lottery) 6 out of 49. ~~Prizes~~:

# correct	Prize
all 6	\$1,200,000
5	\$800
4	\$35

correct	not
6	43

Expected amount the player wins $= X$?

PMF: $p(1{,}200{,}000) = \dfrac{1}{\binom{49}{6}}$

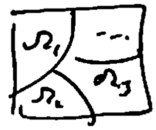
$p(35) = P(4 \text{ correct}, 2 \text{ not}) = \dfrac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}}$

$p(800) = P(5 \text{ correct}, 1 \text{ not}) = \dfrac{\binom{6}{5}\binom{43}{1}}{\binom{49}{6}}$

$$\mathbb{E}[X] = 1{,}200{,}000 \cdot p(1{,}200{,}000) + 800 \cdot p(800) + 35 \cdot p(35) + 0 \cdot p(0) = 0.13$$

Ans: 13¢

Ex (Repeated values) ($\sim$ Remark 2.26)



~~Suppose~~ let $\Omega = \bigcup_k \Omega_k$ be a countable partition of a sample space. Assume that X is constant on each Ω_k, say $X = x_k$ on Ω_k.

possibly the same for different Ω_i

Then $$\mathbb{E}[X] = \sum_k x_k \, \mathbb{P}(\Omega_k).$$

If X takes different values x_k on each Ω_k $\Rightarrow$ QED by def of $\mathbb{E}(X)$

If not, combine the parts Ω_k on which X takes the same value into a bigger set.

Prop (Linearity). $\forall$ r.v. X, $\forall$ const $a \in \mathbb{R}$:

(i) $\mathbb{E}[X+Y] = \mathbb{E}(X) + \mathbb{E}(Y)$;

(ii) $\mathbb{E}(aX) = a \cdot \mathbb{E}(X)$.

(i) $\exists$ partition $\Omega = \bigcup_{ij} \Omega_{ij}$ s.t. both X, Y are const on each Ω_i

($\Omega_{ij} := \{X = x_i, Y = y_j\}$) (Ex.)

$$\Rightarrow \mathbb{E}(X+Y) = \sum_{ij} (x_i + y_j)\, \mathbb{P}(\Omega_{ij})$$

$$= \sum_i x_i \underbrace{\sum_j \mathbb{P}(\Omega_{ij})}_{\mathbb{P}\{X = x_i\}} + \sum_j y_j \underbrace{\sum_i \mathbb{P}(\Omega_{ij})}_{\mathbb{P}\{Y = y_j\}}$$

$$= \mathbb{E}(X) + \mathbb{E}(Y) \quad \text{(by def.)}$$

Remark: (1) Absolute continuity was used here.

(2) For more than 2: $\boxed{\mathbb{E}\left[\sum_i X_i\right] = \sum_i \mathbb{E}[X_i].}$

Other properties (see book):

- $X \le Y \Rightarrow \mathbb{E}(X) \le \mathbb{E}(Y)$
- $X = c$ (const) $\Rightarrow \mathbb{E}[X] = c$
- $a \le X \le b \Rightarrow a \le \mathbb{E}(X) \le b$.
- $f: \mathbb{R} \to \mathbb{R} \Rightarrow \mathbb{E} f(X) = \sum f(x_i)\, \mathbb{P}\{X_i = x_i\}$ (follows from Ex)
- $\mathbb{E}[\mathbb{1}_E] = \mathbb{P}(E)$

indicator

Ex (Matching Problem)

N hw returned to N students at random.

$\mathbb{E}(\underbrace{\#\text{ students getting their own hw}}_{=:X}) = ?$

(Recall that we calculated ~~P{at least 1 stud~~
$\mathbb{P}\{X \geq 1\} = 1 - \frac{1}{e}$).

Method of indicators:

~~X = # stude~~ $X = \sum_1^N X_i$ where $X_i = \begin{cases} 1, & \text{student } i \text{ gets own hw} \\ 0, & \text{———} \end{cases} = \mathbb{1}_{\{i \text{ gets own hw}\}}$

$\Rightarrow \mathbb{E}[X] = \sum_1^N \mathbb{E}[X_i]$.

$\mathbb{E}[X_i] = \mathbb{P}\{\text{Student } i \text{ gets own hw}\} = \frac{1}{N}$.

$\Rightarrow \mathbb{E}(X) = N \cdot \frac{1}{N} = 1$.

~~**Property**: If $X_1, X_2, \ldots$ are random variables, then~~

~~$$E\left[\sum_i X_i\right] = \sum_i E(X_i)$$~~

~~(To be proved later)~~

Example (Group Testing)

Need to test blood of n people for a rare disease (syphilis in men drafted during WWII) $p = P(\text{positive})$.

Method 1: test all people individually ($\Rightarrow$ n tests)

Method 2: draw and mix blood of k people.
(Group testing)
If test is negative $\Rightarrow$ no more tests needed ($\Rightarrow$ 1 test)
If positive $\Rightarrow$ test all k people individually ($\Rightarrow$ $k+1$ tests).
Repeat for next k people, etc.

Expected # of tests in Method 2? $\underbrace{\hspace{6em}}_{X}$

$X_i :=$ # tests on group i.

~~Divide n people into n/k groups of k~~:

$$\underbrace{\underset{X_1}{k}\ \Big|\ \underset{X_2}{k}\ \Big|\ \underset{X_3}{k}\ \Big|\ k\ \Big|\ k\ \Big|\ k\ \Big|\ \underset{X_{n/k}}{k}}_{n/k \text{ groups}}$$

$\Rightarrow$ $X = X_1 + X_2 + \ldots + X_{n/k}$

~~X_i := # of tests needed for group i~~ X_i takes values $1, k+1$,

$P\{X_i = 1\} = P\{\text{all } k \text{ people are negative}\} = (1-p)^k$

$P\{X_i = k+1\} = 1 - (1-p)^k$.

$E[X_i] = 1 \cdot (1-p)^k + (k+1)\left[1 - (1-p)^k\right]$. (same for each group i)

$$\Rightarrow E[X] = \sum_{i=1}^{n/k} E[X_i] = \boxed{\frac{n}{k}\left((1-p)^k + (k+1)\left[1-(1-p)^k\right]\right)}$$

For example, if $n = 100{,}000$, $p = 10^{-4}$ (on ave, 10 sick)

With $k = 120$ $\Rightarrow$ $E[X] = 2026$

Compare to method 1 where # = 100,000.