

MATH 180A: INTRO TO PROBABILITY (FOR DATA SCIENCE)

www.math.ucsd.edu/~tkemp/180A

Today: § 3.3-3.4 Lab 3.2 due **TONIGHT!**

Next: Review (practice exams)

* Regrade requests for HW2: **Tuesday, 10/22**
for Lab 2: **Thursday, 10/24** } **8am - 11pm**

* **Midterm 1**: Wednesday, 10/23, 8-10pm, PCYNH 109
Assigned seats (see Triton Ed)
1 double sided sheet of (hand-written) notes allowed
Practice Midterms posted on webpage

Expectation

Definition: Let X be a discrete random variable with possible values t_1, t_2, t_3, \dots . The **expectation** or **expected value** of X is

$$\mathbb{E}(X) := \sum_j t_j \mathbb{P}(X=t_j)$$

It is also called the **mean** of X , and is often denoted μ .

E.g. Let X be a discrete random variable with probability mass fn.

$$P_X(k) = \frac{c}{k(k+1)}, \quad k = 1, 2, 3, \dots$$

Find c , and compute $\mathbb{E}(X)$.

Expectation of Continuous Random Variable (with density)

X discrete, $X \in \{t_1, t_2, t_3, \dots\}$

$$\mathbb{E}(X) = \sum_j t_j P(X = t_j)$$

X continuous ($P(X=t)=0$
for each $t \in \mathbb{R}$)
↑
probability density $f_X(t)$

Eg. Let $X \sim \text{Unif}([a, b])$. $f_X(t) = \begin{cases} \frac{1}{b-a} & 0 \leq t \leq b \\ 0 & \text{otherwise} \end{cases}$

Question:

Shoot an arrow at a circular target of radius 1. What is the **expected distance** of the arrow from the center?

(a) **1**

(b) **$\frac{2}{3}$**

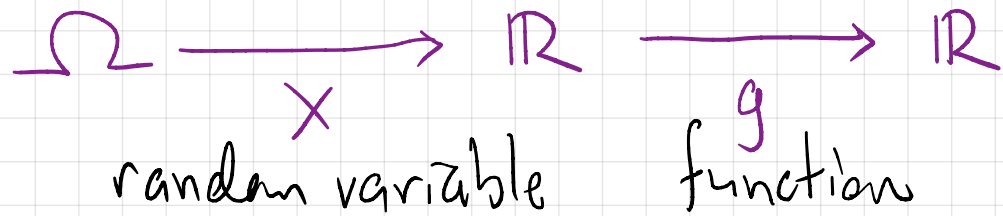
(c) **$\frac{1}{2}$**

(d) **$\frac{1}{4}$**

E.g. $f(t) = \begin{cases} 0 & t \leq 1 \\ 1/t^2 & t > 1 \end{cases}$ is a probability density.

If X is a continuous random variable with $f_X = f$, what is $\mathbb{E}(X)$?

Expectations of Functions of Random Variables



$g(X) = g \circ X$ is a new random variable.

Eg. $X \sim \text{Bin}(n, p)$ is the number of successes in n trials
 $g(X) = X/n$ is the proportion of successful trials.

In general: for a discrete random variable

Proposition: $E(g(X)) = \sum_s g(s) P(X=s)$.

Note, by definition, $E(g(X)) = \sum_t t \cdot P(g(X)=t)$.

Proposition: for a discrete random variable

$$\mathbb{E}(g(X)) = \sum_s g(s) \mathbb{P}(X=s).$$

Pf. We know $\sum_t t \cdot \mathbb{P}(g(X)=t)$

Proposition: For a continuous random variable X with probability density f_X ,

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(t) f_X(t) dt.$$

Eg. Let U be a uniform random variable on $[a, b]$. Then

$$\mathbb{E}(U^2) =$$

E.g. Recall the car accident / insurance example.

An accident causes $\$Y$ of damage to your car, where

Your insurance deductible is $\$500$.

What is the expected amount you pay?

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

X = amount you pay
(neither continuous
nor discrete r.v.)

Variance

3.4

Definition: The **variance** of a random variable X is

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2].$$

I.e., first compute $\mu := \mathbb{E}(X)$. Then apply the function $g(x) = (x - \mu)^2$ to X , and compute $\mathbb{E}(g(X))$.

* If X is discrete, $\text{Var}(X) =$

* If X is continuous, $\text{Var}(X) =$

In any case, $\text{Var}(X) \geq 0$. Its square root is

$\sigma(X) = \sqrt{\text{Var}(X)}$ is called standard deviation.

E.g. If $X \sim \text{Ber}(p)$, $\mathbb{E}(X) = p$, \therefore

E.g. If $U \sim \text{Unif}([a, b])$, $\mathbb{E}(U) = \frac{a+b}{2}$, \therefore

Variance is a measure of how "spread out from the mean" the distribution is. For example:

Theorem: Let X be a random variable with finite expectation $E(X) = \mu$. Then

$$\text{Var}(X) = 0 \quad \text{iff} \quad P(X = \mu) = 1 \\ \text{(i.e. } X \equiv \mu \text{.)}$$

Pf. (\Leftarrow)

(\Rightarrow) [Here we will assume X is discrete, for now.]