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 } 8am - 11pm
  for Lab2: Thursday, 10/24 }

* Midterm 1: Wednesday, 10/23, 8-10pm, PCYNH 109
  Assigned seats (see TritonEd)
  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 $t_1, t_2, t_3, \ldots$ 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.

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

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

Expectation of Continuous Random Variable (with density)

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

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

X continuous $\left(\mathbb{P}(X=t) = 0 \text{ for each } t \in \mathbb{R} \right)$

↑

propability density $f_X(t)$

Eg. Let $X \sim \mathrm{Unif}([a,b])$.

$$f_X(t) = \begin{cases} \frac{1}{b-a} & a \le t \le b \\ 0 & \text{otherwise} \end{cases}$$

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) $2/3$

(c) $1/2$

(d) $1/4$

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

$$\Omega \xrightarrow[X]{} \mathbb{R} \xrightarrow[g]{} \mathbb{R}$$

random variable   function

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

Eg. $X \sim \mathrm{Bin}(n,p)$ 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: $\mathbb{E}(g(X)) = \sum_s g(s)\,\mathbb{P}(X=s)$.

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

Proposition: for a discrete random variable

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

Pf. We know

$$\mathbb{E}(g(X)) = \sum_t t \cdot \mathbb{P}(g(X)=t)$$

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

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

Eg. Recall the car accident / insurance example.
An accident causes $\$Y$ of damage to your car, where
Your insurance deductible is $\$500$. $Y \sim \text{Unif}([100, 1500])$.
What is the expected amount you pay?

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

Variance

3.4

Definition: The variance of a random variable X is

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

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, $\mathrm{Var}(X) =$

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

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

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

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

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

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 $\mathbb{E}(X)=\mu$. Then

$$\operatorname{Var}(X)=0 \quad \text{iff} \quad \mathbb{P}(X=\mu)=1.$$

(i.e. $X \equiv \mu$.)

Pf. ($\Longleftarrow$)

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