APMA 1650 - Spring 2021
Lecture 11

Friday, Feb 12, 2021

Announcements

- Midterm Exam 1 Moved to Monday 2/22
- There will be a Homework 4 due 2/19
- TAs will hold review sessions next week (week 5)

Expectation

From last lecture we introduced the concept of the *expected value* of a discrete random variable.

Definition: Let X be a discrete random variable with range $R_X = \{x_1, x_2, x_3, \ldots\}$. The expected value of X denoted by $EX = E(X) = E[X]$ is defined by

$$EX = \mu_X = \sum_{x_k \in R_X} x_k P_X(x_k)$$

$P_x(x_k) \approx \frac{N_k}{N}$

You should think of the expected value as the *average* or mean value that the random variable takes. μ_X is typically called the mean. It is just a real (deterministic) number! (not random)

Expectation:

Example: $X \sim \text{Bernoulli}(p)$. What is EX?

$$R_X = \{0, 1\}$$

$$\mathbb{E}X = \underset{0}{\cancel{0\,(1-p)}} + 1\,p = p$$

Expectation:

Example: $X \sim \text{Geometric}(p)$. What is EX?

$R_X = \{1, 2, \dots\}$ $\qquad P_X(k) = \overset{\overset{q}{\parallel}}{(1-p)}^{k-1} p$

$k = 1, 2, 3, \dots$

$$EX = \sum_{x_k \in R_X} x_k P_X(x_k) = \sum_{k=1}^{\infty} k q^{k-1} p = p \sum_{k=1}^{\infty} k q^{k-1}$$

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}, \quad \frac{d}{dq} \sum_{k=0}^{\infty} q^k = \sum_{k=0}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2}$$

$$\Rightarrow \quad EX = p \frac{1}{(1-q)^2} = \frac{p}{p^2} = \boxed{1/p}.$$

Expectation:

Example: $X \sim$ Poisson(λ). What is EX?

$$R_X = \{0, 1, 2, \ldots\}, \quad P_X(k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

$$EX = \sum_{x_k \in R_X} x_k P_X(x_k) = \sum_{k=0}^{\infty} \frac{\cancel{k}\, e^{-\lambda} \lambda^k}{\cancel{k}!}$$

$$\lambda^k = \lambda \lambda^{k-1}$$

$$= e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$$

$$= \lambda e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{= e^{\lambda}} = \lambda \cancel{e^{-\lambda}} \cancel{e^{\lambda}} = \boxed{\lambda}.$$

Functions of random variables

Supose X is a discrete random variable and $g : \mathbb{R} \to \mathbb{R}$, then $Y = g(X)$ is another random variable with range

$$R_Y = \{g(x) \, : \, x \in R_X\} = \{g(x_1), g(x_2), \ldots\}.$$

What is the PMF of $Y = g(X)$?

$$P_Y(y) = P(g(X) = y) = \sum_{\{x \,:\, g(x)=y\}} P_X(x)$$

Sum over all the probabilities of values $x \in R_X$ that lead to $g(x) = y$, can have multiple values if g is not one-to-one.

Functions of random variables

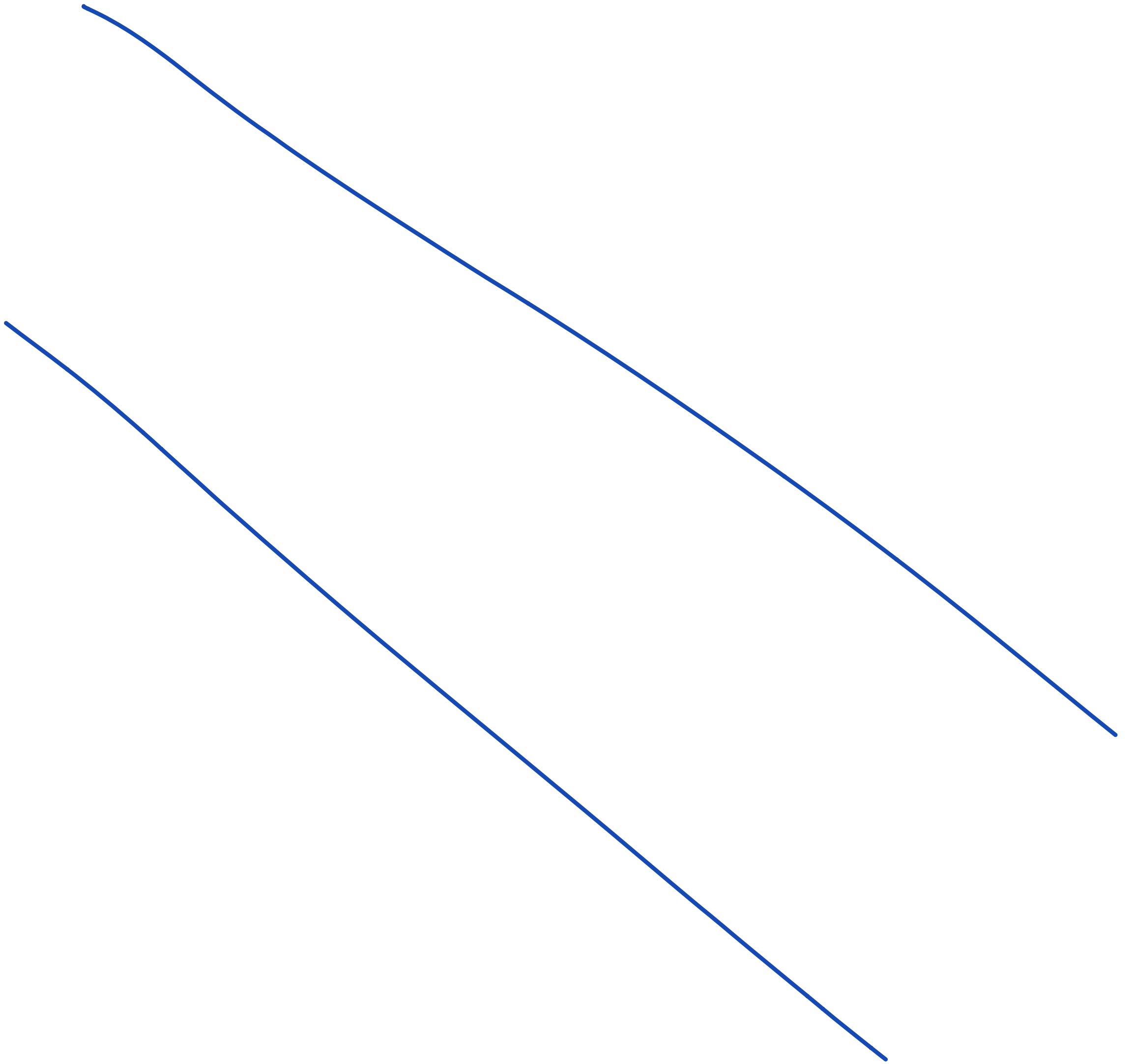
Example X is a discrete random variable with $P_X(k) = \frac{1}{5}$ for $k \in \{-2, -1, 0, 1, 2\}$.

- What is $P_Y(y)$ for $Y = |X|$?
- What is $E|X|$?

$Y = \|X\|$	0	1	2
$P_Y(y)$	$1/5$	$2/5$	$2/5$

$$E|X| = 0 \cdot \tfrac{1}{5} + 1 \cdot \tfrac{2}{5} + 2 \cdot \tfrac{2}{5} = \frac{6}{5}$$

$$1 \times \tfrac{1}{5} + |-1| \times \tfrac{1}{5} \qquad 2 \cdot \tfrac{1}{5} + |-2| \cdot \tfrac{1}{5}$$



Functions of Random Variables

We can calculate $E[g(X)]$ without having to know the PMF of $g(X)$

Law of the unconcious statistician (LOTUS) Let X be a discete RV and let $g : \mathbb{R} \to \mathbb{R}$,then

$$E[g(X)] = \sum_{x_k \in R_X} g(x_k) P_X(x_k)$$

This law is "obvious" to a statistician. Think in terms of relative frequencies...

$$\text{average of } g(X) = \frac{1}{N} \sum_{x_k \in R_X} \{\# \text{ of times } x_k \text{ occurs}\} g(x_k)$$

Proof

$$
\begin{aligned}
E[Y] &= \sum_{y \in R_Y} y P_Y(y) \\
&= \sum_{y \in R_Y} y \sum_{x \,:\, g(x)=y} P_X(x) \\
&= \sum_{y \in \mathbb{R}_X} \sum_{x:g(x)=y} g(x) P_X(x) \\
&= \sum_{x \in R_X} g(x) P_X(x)
\end{aligned}
$$

Here we used the partition

$$
R_X = \bigcup_{y \in R_Y} \{x \,:\, g(x) = y\}
$$

Properties of Expectation

Linearity of Expectation: Let $a, b \in \mathbb{R}$ and X be a discrete RV

Y

$$E[aX + b] = aE[X] + b$$

$$\mu_Y = a\mu_X + b$$

Proof: Take $g(x) = ax + b$, then

$$\begin{aligned} E[g(X)] &= \sum_{x \in R_X} (ax + b) P_X(x) \\ &= a \underbrace{\sum_{x \in R_X} x P_X(x)}_{EX} + b \underbrace{\sum_{x \in R_X} P_X(x)}_{=1} \end{aligned}$$

Properties of Expectation

More Generally: Let $X_1, \ldots, X_n$ be ANY discrete random variables

$$E[X_1 + X_2 + \ldots + X_n] = EX_1 + EX_2 + \ldots + EX_n$$

Note there is no assumption of independence here

We need the concept of joint probability to prove this (this will be later)

Expectation:

Example: $X \sim$ Binomial(n, p). What is EX?

$$EX = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = ?$$

$$X = X_1 + X_2 + \dots + X_n$$

$$\rightsquigarrow X_i \sim \text{Bernoulli}(p)$$

$$EX = \sum_{k=0}^{n} EX_i = \sum_{k=0}^{n} p = \underline{\underline{np}}.$$

Geometric(p).

How about $X \sim$ Pascal(m, p)?

$$X = X_1 + X_2 + \dots + X_m$$

Expectation

Example: n students turn in homework to m dropboxes. Each student picks a dropbox at random. What is the expected number of empty dropboxes?

$$X_i = \begin{cases} 1 & \text{if box } i \text{ is empty} \\ 0 & \text{otherwise} \end{cases}$$

$$X = \sum_{i=1}^{m} X_i = \#\text{ of empty dropboxes.}$$

$$\mathbb{E} X_i = \text{prob box } i \text{ empty}. = \left(\frac{m-1}{m}\right)^n$$

$$\mathbb{E} X = \sum_{i=1}^{m} \mathbb{E} X_i = m\left(\frac{m-1}{m}\right)^n$$

Variance

A measure of how spread out a random variable (or PMF) is about it's mean $\mu_X = EX$

$$E[\overbrace{X - \mu_X}^{\text{how far from mean}}] = E[X] - \mu_X = \mu_X - \mu_X = 0.$$

Definition Let X be a random variable with mean $\mu_X = EX$, then the variance of X is defined by

$$\mathsf{Var}(X) = E[(X - \mu_X)^2]$$

The standard deviation is defined by

$$\mathsf{SD}_X = \sigma_X = \sqrt{\mathsf{Var}(X)}$$

A useful formula

Computational formula: Let X be a random variable with mean $\mu_X = EX$,

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

Proof:

$$\begin{aligned}
\text{Var}(X) &= E(X - \mu_X)^2 \\
&= E[X^2 - 2\mu_X X + \mu_X^2] \\
&= E[X^2] - 2\mu_X \underset{''\mu_X}{EX} + \mu_X^2 \\
&= E[X^2] - 2\mu_X^2 + \mu_X^2 \\
&= E[X^2] - \mu_X^2.
\end{aligned}$$

Variance

Example: $X \sim \text{Bernoulli}(p)$. What is $\text{Var}(X)$?

$$\text{Var}(X) = EX^2 - \mu_X^2$$

$$= \cancel{0^2(1-p)} + 1^2 p - p^2$$

$$= p^2 - p = p(1-p) = pq.$$

Properties of Variance

Rescaling: Let X be a random variable and $a, b \in \mathbb{R}$, then

$$\text{Var}(aX + b) = a^2\text{Var}(X)$$

Proof: Let $Y = aX + b$. Note that $\mu_Y = a\mu_X + b$, so

$$Y - \mu_Y = (aX + b) - (a\mu_X + b) = a(X - \mu_X)$$

Therefore

$$\text{Var}(Y) = E[(Y - \mu_Y)^2] = E[a^2(X - \mu_X)^2] = a^2\text{Var}(X)$$

Shifting a random variable by a constant doesn't change it's variance. Rescaling

Properties of Variance

More generally: Let $X_1, X_2, \ldots X_n$ be independent random variables, then

$$\boxed{\text{Var}(X_1 + X_2 + \ldots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \ldots \text{Var}(X_n).}$$

We will be able to show this later using the idea of covariance.

This is not true if the random variables are not independent!

Variance

Example: $X \sim$ Binomial(n, p). What is Var(X)?

$X = X_1 + X_2 + \dots + X_n$, $X_i \sim$ Bernoulli(p)

independent.

$$\mathrm{Var}(X) = \sum_{k=1}^{n} \underbrace{\mathrm{Var}(X_k)}_{= p(1-p)} = n\,p(1-p).$$