10-2: Variance
CSCI 3022 Fall 19

What is the **expected value** of the Bernoulli distribution with parameter p?

Opening sol:

The pmf of the Bernoulli is given by

$$f(x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$$

$$E[X] = \sum_{x \in \Omega} x f(x)$$

$$= 0 \cdot (1-p) + 1 \cdot p$$

$$= p.$$

Opening sol:

The pmf of the Bernoulli is given by

$$f(x) = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \end{cases}$$

We now must sum over both outcomes while multiplying by the probability of those outcomes:

$$\sum_{x \in \{0,1\}} x\, f(x) = \sum_{x \in \{0,1\}} x\, P(X = x)$$

$$= 0 \cdot P(X = 0) + 1 \cdot P(X = 1) = 0 \cdot (1 - p) + 1 \cdot p = \boxed{p}$$

... which makes perfect sense, since p is how often we "expect" a heads.

Announcements and Reminders

- quizlet04 for Friday.
- Practicum 1 posted today.
- Midterm 1; Tuesday, October, 8.
- 6:30-8:00 PM on Tuesday 8 October, this section in HUMN 1B50: Inform me as soon as possible about accommodations.

EV Recap

Population: greek letter

Sample: english

1. **Expected Value:** The average value for X coming from a distribution (not a sample!). Denoted $E[X]$ or μ or μ_X.

   Discrete: $\sum_{x\in\Omega} xf(x)$; Continuous: $\int_{x\in\Omega} \boxed{x} \cdot \boxed{f(x)}\, dx$

   μ $\tilde{\ }$ $\bar{X}$

2. Expected value of a function $g(X)$ of X is:

   $\sum_{x\in\Omega} g(x)f(x)$; $\int_{x\in\Omega} \underbrace{g(x)}_{\text{outcome}} \cdot \underbrace{f(x)}_{\text{probability}}\, dx$

3. $Y = g(X)$ is a *change of variables*.

4. Expectation is **linear:** $E[aX+b] = aE[X] + b$ Proof:

$E[aX+b] = \sum (ax+b)\cdot f(x) = \sum [axf(x) + bf(x)]$

courses

$= a\sum xf(x) + b\sum f(x) = a\sum xf(x) + b$

EV Recap

1. **Expected Value:** The average value for X coming from a distribution (not a sample!). Denoted $E[X]$ or μ or μ_X.
   Discrete: $\displaystyle\sum_{x\in\Omega} xf(x)$; Continuous: $\displaystyle\int_{x\in\Omega} x \cdot f(x)\,dx$
2. Expected value of a function $g(X)$ of X is:
   $\displaystyle\sum_{x\in\Omega} g(x)f(x)$; $\displaystyle\int_{x\in\Omega} g(x) \cdot f(x)\,dx$
3. $Y = g(X)$ is a *change of variables*.
4. Expectation is **linear:** $E[aX+b] = aE[X]+b$ Proof: $E[aX+b] = \int (ax+b)f(x)\,dx = a\int xf(x)\,dx + b\int f(x)\,dx = aE[X]+b$, since integration is also linear!

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.)

https://www.youtube.com/watch?v=naUppHrHJpI

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What *distribution* does X follow?

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What *distribution* does X follow?

> Each row results in a move right with probability p, and a move left with probability $1-p$. We have n rows... or *trials*

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What *distribution* does X follow?

Each row results in a move right with probability p, and a move left with probability $1-p$. We have n rows... or *trials*

Each row is a *Bernoulli*, and our ending bucket is the total number of right-hand moves over the entire experiment, or the sum of n Bernoullis!

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What *distribution* does X follow?

Each row results in a move right with probability p, and a move left with probability $1-p$. We have n rows... or *trials*

independent
identical distribution

For $Y_i \overset{iid}{\sim} Bern(p)$, we have $X = \sum Y_i$. So X is a **binomial** with parameters n and p.

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *expected value* of X?

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *expected value* of X?

> We can use our knowledge about Bernoullis! X is a sum of Bernoulli r.v.s that each have mean or expected value of p.

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *expected value* of X?

We can use our knowledge about Bernoullis! X is a sum of Bernoulli r.v.s that each have mean or expected value of p.

$E[X] = E[Y_1 + Y_2 + Y_3 + \dots Y_n]$, then use linearity:

row1 row2 row$_n$

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *expected value* of X?

We can use our knowledge about Bernoullis! X is a sum of Bernoulli r.v.s that each have mean or expected value of p.

$E[X] = E[Y_1 + Y_2 + Y_3 + \dots Y_n]$, then use linearity:
$E[X] = E[Y_1] + E[Y_2] + E[Y_3] + \cdots + E[Y_n]$. This works even though each Y_i is also a random variable!

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *expected value* of X?

We can use our knowledge about Bernoullis! X is a sum of Bernoulli r.v.s that each have mean or expected value of p.

$E[X] = E[Y_1 + Y_2 + Y_3 + \dots Y_n]$, then use linearity:
$E[X] = E[Y_1] + E[Y_2] + E[Y_3] + \cdots + E[Y_n]$. This works even though each Y_i is also a random variable!

$E[X] = p + p + p + \cdots + p = np$, since each Y is identical.
... which again makes perfect sense, since it's n tries that have a per-try expected value of p. also: $\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \cdot x$

$= np$

Plinko... is random?

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *Variance* of X?

Plinko... is random?

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *Variance* of X?

Recall: Sample *Variance* is $\dfrac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}$

Plinko... is random?

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *Variance* of X?

Another way: sample variance is $\underbrace{\frac{1}{n-1}\sum_{i=1}^{n}}_{\text{averaged out}} \underbrace{\left(X_i - \bar{X}\right)^2}_{\text{squared deviations}}$

Population variance is this idea expressed as an expectation:

Plinko... is random?

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *Variance* of X?

Another way: sample variance is $\underbrace{\frac{1}{n-1}\sum_{i=1}^{n}}_{\text{averaged out}} \underbrace{\left(X_i - \bar{X}\right)^2}_{\text{squared deviations}}$

Population variance is this idea expressed as an expectation:

$\bar{X} \to \mu$

Sample $\to$ Pops.

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

E: "average" for a pop

Variance of a Random Variable

Compare to s^2?

Definition: *Variance:*

For a discrete random variable X with pdf $f(x)$, the *variance* of X is denoted as $Var[X] = \sigma^2$ and is calculated as:

$$Var[X] = E[(X - E[X])^2]$$

1. Continuous: $\underbrace{\int_{x \in \Omega}}$ $\underbrace{\text{outcome}}$ $\cdot$ Prob $= \int (x-\mu_x)^2 f(x)\,dx$

2. Discrete:

outcome X mean of X.

$$= \sum_{x \in \Omega} (x - E[X])^2 f(x)$$

pop

The standard deviation (SD) of X is: $\sigma = \sqrt{\sigma^2}$: Units are X-units.

Variance of a Random Variable

Definition: *Variance:*

For a discrete random variable X with pdf $f(x)$, the *variance* of X is denoted as $\underline{Var[X] = \sigma^2}$ and is calculated as:

$$Var[X] = E[(X - E[X])^2]$$

1. Continuous:

$$Var[X] = \int_{x \in \Omega} (x - \mu_x)^2 \cdot f(x)\, dx$$

2. Discrete:

$$Var[X] = \sum_{x \in \Omega} (x - \mu_x)^2 f(x)$$

The standard deviation (SD) of X is: $\sigma = \sqrt{\sigma^2}$

Variance Calculated

We want more Plinko! Let's find the variance of a Bernoulli so we can build on it.

Recall: The pmf of the Bernoulli is given by

$$f(x) = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \end{cases}$$

and we know that $\underline{E[X] = p}$.

$$\sum_{x \in \{0,1\}} (x-p)^2 f(x) = (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p$$

$$= p^2(1-p) + p(1-p)^2$$

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

Variance Calculated

We want more Plinko! Let's find the variance of a Bernoulli so we can build on it.

Recall: The pmf of the Bernoulli is given by

$$f(x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$$

and we know that $E[X] = p$. We now must sum over both outcomes' deviations from the mean while multiplying by those probabilities

$$E[(X - E[X])^2] = \sum_{x \in \{0,1\}} (x-p)^2 f(x) = \sum_{x \in \{0,1\}} (x-p)^2 P(X = x)$$

$$= (0-p)^2 \cdot P(X=0) + (1-p)^2 \cdot P(X=1) = (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p$$

$$= (p)(1-p)(p+1-p) = \boxed{p(1-p)}$$

Let's play Plinko!!

Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *variance* of X follow?

Need to know: if two random variables are **independent**,

$$\boxed{Var[X+Y] = Var[X] + Var[Y]}$$

Really, really important

Usually false

Let's play Plinko!!

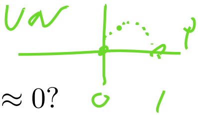
Let X be the random variable describing the result in each round of Plinko with n rows and probability p of moving to the right off of each peg. (Ignoring the edges for now.) What is the *variance* of X follow?

Need to know: if two random variables are **independent**,

$$\boxed{Var[X+Y] = Var[X] + Var[Y]}$$

So for Plinko, where $X = Y_1 + Y_2 + \cdots + Y_i$ but the $Y_i's$ are all independent,

$$Var[X] = Var[\sum Y_i] \underbrace{=}_{\text{indep}} \sum Var[Y] \underbrace{=}_{\text{ident}} nVar[Y_i] = \boxed{np(1-p)}$$



Sanity Check! Should variance be smaller if $p \approx 1$ or $p \approx 0$?

Let's talk Variance

For a random variable X and constants a and b, if we define $Y = aX + b$...

$E[Y] = aE[X] + b$ because Expectation $E[\cdot]$ is **linear**. Is $Var[\cdot]$?

1. What is $Var[X + b]$?

   Probably $Var[X]$; dispersion didn't change.

2. What is $Var[aX]$?

   spreads/contracts

Let's talk Variance

For a random variable X and constants a and b, if we define $Y = aX + b$...
$E[Y] = aE[X] + b$ because Expectation $E[\cdot]$ is **linear**. Is $Var[\cdot]$?

1. What is $Var[X + b]$?
   Intuition: moving X doesn't change its spread!
2. What is $Var[aX]$?
   Intuition: multiplying X should change its spread!

Non-linear Variance

$Var[X] = E[(X - E(X))^2]$

For a random variable X and constants a and b, if we define $Y = aX + b$...

What is $Var[aX + b]$? $= \int \left((ax+b) - \underbrace{E[ax+b]}_{aE[X]+b}\right)^2 f(x)\,dx$

$= \int [ax + b - aE[x] - b]^2 f(x)\,dx$

$= \int [a(x - E[x])]^2 f(x)\,dx$

$= \int a^2 (x - E[x])^2 f(x)\,dx = a^2 Var[x]$

Non-linear Variance

For a random variable X and constants a and b, if we define $Y = aX + b$...

What is $Var[aX + b]$?

$$\begin{aligned} Var[aX+b] &= \sum_{x\in\Omega}(aX + b - E[aX+b])^2 f(x) \\ &= \sum_{x\in\Omega}(aX + b - a\overbrace{E[X]}^{\mu_X} - b)^2 f(x) \\ &= \sum_{x\in\Omega}(aX - aE[X])^2 f(x) \\ &= \sum_{x\in\Omega} a^2 (X - E[X])^2 f(x) \\ &= a^2 \sum_{x\in\Omega}(X - E[X])^2 f(x) \\ &= \boxed{a^2 Var[X]} \end{aligned}$$

Really non-linear Variance

$Z = X + X = 2X$

$Z = X - X = 0$

What if we want to know what happens to two events that *aren't* independent? For example, what's the variance of $Z = X + Y$?

Really non-linear Variance

What if we want to know what happens to two events that *aren't* independent? For example, what's the variance of $Z = X + Y$? $Y = f(x)$,

$$Var[X+Y] = \sum_{x \in \Omega} (X + Y - E[X+Y])^2 f(x)$$

$f(x,y)$,

$$Var[X+Y] = \sum_{x \in \Omega} (X + Y - E[X] - E[Y])^2 f(x) \quad f(x,y).$$

Really non-linear Variance

What if we want to know what happens to two events that *aren't* independent? For example, what's the variance of $Z = X + Y$?

$$Var[X+Y] = \sum_{x \in \Omega} (X + Y - E[X+Y])^2 f(x)$$

$$Var[X+Y] = \sum_{x \in \Omega} (X + Y - E[X] - E[Y])^2 f(x)$$

If we expand this out, we have to deal with a bunch of $XY, XE[Y],$ etc. terms. It matters if X and Y move *together*. It helps to define this concept. What does it mean for X and Y to move *together*?
Example: what if $Y = -X$? Then the variance of Z is zero!

Covariance

$E[(X-\mu_X)^2]$

When two random variables X and Y are not independent, it is frequently of interest to assess how strongly they are related to one another.

Definition: *Covariance:*

The covariance between two rv's X and Y is defined as:

++ if $X > \mu_X$, $Y > \mu_Y$

$$E[\underbrace{(X-\mu_X)}_{\text{X versus its mean}} \overbrace{(Y-\mu_Y)}^{\text{Y versus its mean}}]$$

−− if $X < \mu_X$, $Y < \mu_Y$

If both variables tend to deviate in the same direction (both go above their means or below their means at the same time), then the covariance will be positive.

If the opposite is true, the covariance will be negative.

If X and Y are not strongly related, the covariance will be near 0.

Units: X units times Y units

Correlation

Definition: *Correlation*
The *correlation* coefficient of X and Y, denoted by $Cor[X,Y]$ or just ρ, is the *unitless* measure of covariance defined by:

$$Cor[X,Y] = \frac{Cov[X,Y]}{\sigma_X \sigma_Y} = \frac{E[(X-E[X])(Y-E[Y])]}{\sigma_X \sigma_Y}$$

It represents a "scaled" covariance: correlation ranges between -1 and 1.

Correlation

Definition: *Correlation*
The *correlation* coefficient of X and Y, denoted by $\underline{Cov[X, Y]}$ or just $\underline{\rho}$, is the *unitless* measure of covariance defined by:

$$\rho = \frac{Cov[X, Y]}{\sigma_X \sigma_Y}$$

It represents a "scaled" covariance: correlation ranges between -1 and 1.

Covariance Pictured

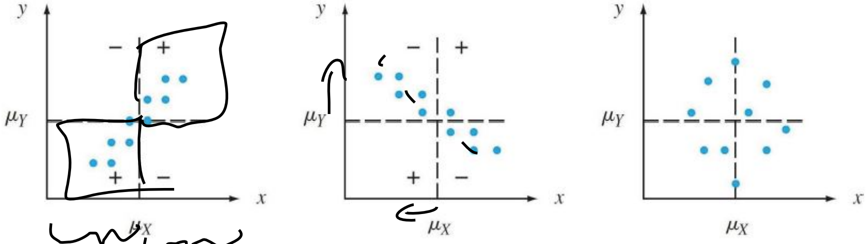
The covariance depends on both the set of possible pairs and the probabilities of those pairs.

Below are examples of 3 types of “co-varying”:

Cov, Cor > 0
"slope" > 0.

Cov, Cor < 0
slope negative

Cor ≈ 0.

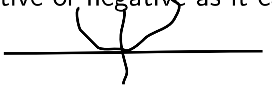


$X < \mu_X$ $\quad$ $X > \mu_X$

Interpreting Correlation

If X and Y are independent, then $\rho=0$ _____, but $\rho=0$ _____ does not imply independence.

The correlation coefficient is a measure of the *linear relationship* between X and Y, and only when the two variables are perfectly related in a *linear* manner will be as positive or negative as it can be.



Two variables could be uncorrelated yet highly dependent because there is a strong nonlinear relationship, so be careful not to conclude too much from low correlation.

We return to covariance in a few weeks...

Interpreting Correlation

If X and Y are independent, then $\underline{\rho = 0}$, but $\underline{\rho = 0}$ does not imply independence.

The correlation coefficient is a measure of the *linear relationship* between X and Y, and only when the two variables are perfectly related in a *linear* manner will be as positive or negative as it can be.

Two variables could be uncorrelated yet highly dependent because there is a strong nonlinear relationship, so be careful not to conclude too much from low correlation.

We return to covariance in a few weeks...

Interpreting Correlation