THE UNIVERSITY OF TEXAS AT AUSTIN
Department of Statistics and Data Sciences
College of Natural Sciences

SDS 321: Introduction to Probability and Statistics
Lecture 13: Expectation and Variance and joint distributions

Purnamrita Sarkar
Department of Statistics and Data Science
The University of Texas at Austin

www.cs.cmu.edu/∼psarkar/teaching

Multiple random variables

So far we have been talking about single random variables and associated PMF's. However, often we are interested in multiple random variables.

- Consider two discrete random variables X, and Y associated with the same experiment.
- The joint PMF of X and Y are defined as $p_{X,Y}(x, y) = P(X = x, Y = y)$ for all pairs of values x, y X and Y can take.
- This is none other than $P(\{X = x\} \cap \{Y = y\})$.
- Of course the order does not matter.

Properties of the joint PMF

- Recall that if $A_1, A_2, \ldots, A_K$ is a partition of Ω,
$$P(B) = P\left(\bigcup_k (B \cap A_k)\right) = \sum_k P(B \cap A_k).$$
- $\{X = x\}$ is the disjoint union of $\{X = x\} \cap \{Y = y\}$ for all y values Y can take.
- $\{X = x\} \cap \{Y = y\}$ is none other than $\{X = x, Y = y\}$
- We can extend this to PMFs: $\sum_y P(X = x, Y = y) =$.
$$\sum_x P(X = x, Y = y) =$$

Properties of the joint PMF

- Recall that if $A_1, A_2, \ldots, A_K$ is a partition of Ω,

$$P(B) = P\left(\bigcup_k (B \cap A_k)\right) = \sum_k P(B \cap A_k).$$

- $\{X = x\}$ is the disjoint union of $\{X = x\} \cap \{Y = y\}$ for all y values Y can take.
- $\{X = x\} \cap \{Y = y\}$ is none other than $\{X = x, Y = y\}$
- We can extend this to PMFs: $\sum_y P(X = x, Y = y) = P(X = x)$.

$$\sum_x P(X = x, Y = y) =$$

Properties of the joint PMF

- Recall that if $A_1, A_2, \ldots, A_K$ is a partition of Ω,
$$P(B) = P\left(\bigcup_k (B \cap A_k)\right) = \sum_k P(B \cap A_k).$$
- $\{X = x\}$ is the disjoint union of $\{X = x\} \cap \{Y = y\}$ for all y values Y can take.
- $\{X = x\} \cap \{Y = y\}$ is none other than $\{X = x, Y = y\}$
- We can extend this to PMFs: $\sum_y P(X = x, Y = y) = P(X = x)$.
$$\sum_x P(X = x, Y = y) = P(Y = y)$$
- These are also called the marginal PMF's.

Properties of the joint PMF

- Recall that if $A_1, A_2, \ldots, A_K$ is a partition of Ω,
$$P(B) = P\left(\bigcup_k (B \cap A_k)\right) = \sum_k P(B \cap A_k).$$
- $\{X = x\}$ is the disjoint union of $\{X = x\} \cap \{Y = y\}$ for all y values Y can take.
- $\{X = x\} \cap \{Y = y\}$ is none other than $\{X = x, Y = y\}$
- We can extend this to PMFs: $\sum_y P(X = x, Y = y) = P(X = x)$.
$$\sum_x P(X = x, Y = y) = P(Y = y)$$
- These are also called the marginal PMF's.
- So $\sum_x \sum_y P(X = x, Y = y) =$.

Properties of the joint PMF

- Recall that if $A_1, A_2, \ldots, A_K$ is a partition of Ω,

$$P(B) = P\left(\bigcup_k (B \cap A_k)\right) = \sum_k P(B \cap A_k).$$

- $\{X = x\}$ is the disjoint union of $\{X = x\} \cap \{Y = y\}$ for all y values Y can take.
- $\{X = x\} \cap \{Y = y\}$ is none other than $\{X = x, Y = y\}$
- We can extend this to PMFs: $\sum_y P(X = x, Y = y) = P(X = x)$.

$$\sum_x P(X = x, Y = y) = P(Y = y)$$

- These are also called the marginal PMF's.

- So $\sum_x \sum_y P(X = x, Y = y) =$.
  - $\{X = x, Y = y\}$ for all pairs of numerical values taken by X and Y form a partition of the sample space.
  - And now the normalization rule gives us the result!

Properties of the joint PMF

- Recall that if $A_1, A_2, \ldots, A_K$ is a partition of Ω,
$$P(B) = P\left(\bigcup_k (B \cap A_k)\right) = \sum_k P(B \cap A_k).$$
- $\{X = x\}$ is the disjoint union of $\{X = x\} \cap \{Y = y\}$ for all y values Y can take.
- $\{X = x\} \cap \{Y = y\}$ is none other than $\{X = x, Y = y\}$
- We can extend this to PMFs: $\sum_y P(X = x, Y = y) = P(X = x)$.
$$\sum_x P(X = x, Y = y) = P(Y = y)$$
- These are also called the marginal PMF's.

- So $\sum_x \sum_y P(X = x, Y = y) = \sum_x P(X = x) = 1$.
  - $\{X = x, Y = y\}$ for all pairs of numerical values taken by X and Y form a partition of the sample space.
  - And now the normalization rule gives us the result!

Contingency tables

Alice says that there are more left handed women than left handed men. Bob gives her some numbers to count probabilities.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(X = 0, L = 0) =$
- $P(X = 0, L = 1) =$
- $P(X = 1, L = 0) =$
- $P(X = 1, L = 1) =$
- $P(X = 1) =$
- $P(L = 1) =$

Contingency tables

Alice says that there are more left handed women than left handed men. Bob gives her some numbers to count probabilities.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(X = 0, L = 0) = 43/100$
- $P(X = 0, L = 1) =$
- $P(X = 1, L = 0) =$
- $P(X = 1, L = 1) =$
- $P(X = 1) =$
- $P(L = 1) =$

Contingency tables

Alice says that there are more left handed women than left handed men. Bob gives her some numbers to count probabilities.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(X = 0, L = 0) = 43/100$
- $P(X = 0, L = 1) = 7/100$
- $P(X = 1, L = 0) =$
- $P(X = 1, L = 1) =$
- $P(X = 1) =$
- $P(L = 1) =$

Contingency tables

Alice says that there are more left handed women than left handed men. Bob gives her some numbers to count probabilities.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(X = 0, L = 0) = 43/100$
- $P(X = 0, L = 1) = 7/100$
- $P(X = 1, L = 0) = 47/100$
- $P(X = 1, L = 1) = 3/100$
- $P(X = 1) =$
- $P(L = 1) =$

Contingency tables

Alice says that there are more left handed women than left handed men. Bob gives her some numbers to count probabilities.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(X = 0, L = 0) = 43/100$
- $P(X = 0, L = 1) = 7/100$
- $P(X = 1, L = 0) = 47/100$
- $P(X = 1, L = 1) = 3/100$
- $P(X = 1) = 1/2 \leftarrow$ Marginal probability!
- $P(L = 1) = 10/100 \leftarrow$ Marginal probability!

Contingency tables

Alice says that there are more left handed women than left handed men. Bob gives her some numbers to count probabilities.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(X = 0, L = 0) = 43/100$
- $P(X = 0, L = 1) = 7/100$
- $P(X = 1, L = 0) = 47/100$
- $P(X = 1, L = 1) = 3/100$
- $P(X = 1) = 1/2 \leftarrow$ Marginal probability!
- $P(L = 1) = 10/100 \leftarrow$ Marginal probability!
- Remember! These really are estimated numbers, and hence approximations. I am estimating the fraction of left handed men in a population via my sample!

Functions of multiple random variables

- $E(g(X,Y)) = \sum_{x,y} g(x,y)P(X=x, Y=y)$.
- Let $g(X,Y) = aX + bY$.
  - $E(g(X,Y)) = \sum_{x,y}(ax+by)P(X=x, Y=y) = aE[X] + bE[Y]$.
- What if $g(X,Y) = aX^2 + bY^2 + c$?
  - $E[g(X,Y)] = aE[X^2] + bE[Y^2] + c$
  - **Common Mistake:** $E[g(X,Y)] \neq g(E[X], E[Y])$! unless g is linear in X and Y!

Multiple random variables

How about three random variables?

- We will write $p_{X,Y,Z}(x, y, z) = P(X = x, Y = y, Z = z)$
- The rules are the same:
  - $P(X = x, Y = y) = \sum_{z} P(X = x, Y = y, Z = z)$.
  - $P(X = x) = \sum_{y,z} P(X = x, Y = y, Z = z)$.
  - $P(Y = y) = \sum_{x,z} P(X = x, Y = y, Z = z)$.
  - $P(Z = z) = \sum_{x,y} P(X = x, Y = y, Z = z)$.
  - $\sum_{x,y,z} P(X = x, Y = y, Z = z) = 1$.
- Generalizes easily to more than 3 random variables.

Linearity of expectation

Perhaps one of the most useful and powerful results!

- $E[aX + bY + cZ + d] = aE[X] + bE[Y] + cE[Z] + d$
- More generally,

  $E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \ldots a_nE[X_n]$
- This is extremely general! $X_1, \ldots, X_n$ do not have to be mutually independent for this to hold!
- This generalizes to $E\left[\sum_i a_i f(X_i)\right] = \sum_i a_i E\left[f(X_i)\right]$, as long as the expectations are well defined.

Expectation of $Y \sim Binomial(n, p)$

Remember that a Binomial(n, p) random variable is nothing other than the sum of n independent Bernoulli's!

- $Y = \sum_{i=1}^{n} X_i$, where $X_i \sim$ Bernoulli(n, p).
- We know that $E[X_i] = p$.
- Using our newfound tool, we have:

  $$E[Y] = E[\sum_i X_i] = \sum_i E[X_i] = np.$$

- We do not need the mutual independence of the Bernoullis to get this result!

Balls and bins

I am throwing m distinguishable balls into n distinguishable bins. What is the expected number of empty bins (call this Y)?

- Let $X_i = \begin{cases} 1 & \text{The } i^{th} \text{ bin is empty} \\ 0 & \text{Otherwise} \end{cases}$
- We want $E[Y]$.
- $E[Y] =$
- $E[X_i] =$

Balls and bins

I am throwing m distinguishable balls into n distinguishable bins. What is the expected number of empty bins (call this Y)?

- Let $X_i = \begin{cases} 1 & \text{The } i^{th} \text{ bin is empty} \\ 0 & \text{Otherwise} \end{cases}$
- We want $E[Y]$.
- $E[Y] = E[\sum_i X_i] = \sum_i E[X_i]$
- $E[X_i] =$

Balls and bins

I am throwing m distinguishable balls into n distinguishable bins. What is the expected number of empty bins (call this Y)?

- Let $X_i = \begin{cases} 1 & \text{The } i^{th} \text{ bin is empty} \\ 0 & \text{Otherwise} \end{cases}$
- We want $E[Y]$.
- $E[Y] = E[\sum_i X_i] = \sum_i E[X_i]$
- $E[X_i] = P(\text{No ball falls in bin } i) =$

Balls and bins

I am throwing m distinguishable balls into n distinguishable bins. What is the expected number of empty bins (call this Y)?

- Let $X_i = \begin{cases} 1 & \text{The } i^{th} \text{ bin is empty} \\ 0 & \text{Otherwise} \end{cases}$
- We want $E[Y]$.
- $E[Y] = E[\sum_i X_i] = \sum_i E[X_i]$
- $E[X_i] = P(\text{No ball falls in bin } i) = (1 - 1/n)^m$
- $E[Y] = n(1 - 1/n)^m$

Balls and bins

I am throwing m distinguishable balls into n distinguishable bins. What is the expected number of empty bins (call this Y)?

- Let $X_i = \begin{cases} 1 & \text{The } i^{th} \text{ bin is empty} \\ 0 & \text{Otherwise} \end{cases}$
- We want $E[Y]$.
- $E[Y] = E[\sum_i X_i] = \sum_i E[X_i]$
- $E[X_i] = P(\text{No ball falls in bin } i) = (1 - 1/n)^m$
- $E[Y] = n(1 - 1/n)^m$
- When $m = n$, for large n, $E[Y] = n(1 - 1/n)^n \approx n/e$.

Conditional PMF

So we have started thinking about how *knowing about one random variable* alters out belief about another random variable. This brings us to conditional PMFs!

- The **conditional PMF** of a random variable X, conditioned on a particular event A with $P(A) > 0$, is defined by:

  $$p_{X|A}(x) = P(X = x|A) = \frac{P(\{X = x\} \cap A)}{P(A)}$$

- So we have
  $$\sum_x P(X = x|A) = \sum_x \frac{P(\{X = x\} \cap A)}{P(A)} = \frac{\sum_x P(\{X = x\} \cap A)}{P(A)}$$
- But A can be written as a disjoint union of the events $\{X = x\} \cap A$ for all numerical values X takes.
- Total probability rule gives: $P(A) = \sum_x P(\{X = x\} \cap A)$, and so

  $$\sum_x P(X = x|A) = 1.$$

Conditioning one random variable on another

Let X and Y be two random variables associated with the same experiment. Now the knowledge of $Y = y$ for some particular value y provides us with partial knowledge about what value X may take.

- The **conditional PMF** of X given Y is given by $p_{X|Y}(x, y) = P(X = x|\{Y = y\})$.
- Using the same set of rules as before we can write:

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

- For any fixed y such that $P(Y = y) > 0$, we also have: $\sum_x P(X = x|Y = y) = 1$.
- So, a conditional PMF satisfies the properties of a PMF.

Conditional PMF

Bob and Alice are interested in finding out the conditional probability of being left handed given a person is a man. Bob finds his data again.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(L = 1|X = 0)$ is just the fraction of all men who are left handed

Conditional PMF

Bob and Alice are interested in finding out the conditional probability of being left handed given a person is a man. Bob finds his data again.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(L = 1|X = 0)$ is just the fraction of all men who are left handed
- $P(L = 1|X = 0) = 7/50$.

Conditional PMF

Bob and Alice are interested in finding out the conditional probability of being left handed given a person is a man. Bob finds his data again.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(L = 1|X = 0)$ is just the fraction of all men who are left handed
- $P(L = 1|X = 0) = 7/50$.
- Let us plug in the formula. $P(L = 1, X = 0) =$.

Conditional PMF

Bob and Alice are interested in finding out the conditional probability of being left handed given a person is a man. Bob finds his data again.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(L = 1|X = 0)$ is just the fraction of all men who are left handed
- $P(L = 1|X = 0) = 7/50$.
- Let us plug in the formula. $P(L = 1, X = 0) = 7/100$.

Conditional PMF

Bob and Alice are interested in finding out the conditional probability of being left handed given a person is a man. Bob finds his data again.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	**43**	**7**	**50**
Women (X=1)	**47**	**3**	**50**
	90	**10**	**100**

- $P(L = 1|X = 0)$ is just the fraction of all men who are left handed
- $P(L = 1|X = 0) = 7/50$.
- Let us plug in the formula. $P(L = 1, X = 0) = 7/100$.
- $P(X = 0) = 50/100$. So $\dfrac{P(L = 1, X = 0)}{P(X = 0)} = 7/50$.

Conditional PMF

Bob and Alice are interested in finding out the conditional probability of being left handed given a person is a man. Bob finds his data again.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(L = 1|X = 0)$ is just the fraction of all men who are left handed
- $P(L = 1|X = 0) = 7/50$.
- Let us plug in the formula. $P(L = 1, X = 0) = 7/100$.
- $P(X = 0) = 50/100$. So $\frac{P(L = 1, X = 0)}{P(X = 0)} = 7/50$.
- What is $P(L = 0|X = 0)$? Its just the fraction of all men who are right handed! So $43/50$.

Conditional PMF

Bob and Alice are interested in finding out the conditional probability of being left handed given a person is a man. Bob finds his data again.

	Right Handed (L=0)	Left handed (L=1)	
Men (X=0)	43	7	50
Women (X=1)	47	3	50
	90	10	100

- $P(L = 1|X = 0)$ is just the fraction of all men who are left handed
- $P(L = 1|X = 0) = 7/50$.
- Let us plug in the formula. $P(L = 1, X = 0) = 7/100$.
- $P(X = 0) = 50/100$. So $\dfrac{P(L = 1, X = 0)}{P(X = 0)} = 7/50$.
- What is $P(L = 0|X = 0)$? Its just the fraction of all men who are right handed! So $43/50$.
- $P(L = 0|X = 0) + P(L = 1|X = 0) = 1$!

Conditional PMF

- Remember that a conditional PMF is a valid PMF.
- Since $P(X = x|Y = y) = \dfrac{P(X = x, Y = y)}{P(Y = y)}$, we also have the multiplication rule:
  - $P(X = x, Y = y) = P(X = x|Y = y)P(Y = y)$
  - But $P(X = x, Y = y) = P(Y = y, X = x)$, and so we also have: $P(X = x, Y = y) = P(Y = y|X = x)P(X = x)$.
- Same as multiplication rule from before!
- We can also draw trees to get conditional probabilities!

Independence of random variables

- Lets first consider two events $\{X = x\}$ and A. We know that these two events are independent if $P(\{X = x\}, A) = P(\{X = x\})P(A)$
- In other words if $P(A) > 0$, then $P(X = x|A) = P(X = x)$, i.e. knowing the occurrence of A does not change our belief about $\{X = x\}$.
- We will call the random variable X and event A to be independent if

$$P(X = x, A) = P(X = x)P(A) \qquad \textbf{For all } x$$

- Two random variables are said to be independent if

$$P(X = x, Y = y) = P(X = x)P(Y = y) \qquad \textbf{For all } x \textbf{ and } y$$

- To put it a bit differently,

$$P(X = x|Y = y) = P(X = x) \qquad \textbf{For all } x \textbf{ and } y \textbf{ such that } P(Y = y) > 0$$

A super important implication

We saw that $E[X + Y] = E[X] + E[Y]$ no matter whether X and Y are independent or not.

- If X and Y are independent, $E[XY] = E[X]E[Y]$

- $$E[XY] = \sum_{x,y} xyP(X = x, Y = y) = \sum_{x,y} xyP(X = x)P(Y = y)$$
$$= \left(\sum_x xP(X = x)\right)\left(\sum_y yP(Y = y)\right) = E[X]E[Y]$$

- In fact, $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$

Variance of sum of independent random variables

Let X and Y be two independent random variables. What is $\mathrm{var}(X+Y)$?

- Remember! $\mathrm{var}(X+Y) = E[(X+Y)^2] - (E[X+Y])^2$
- $$\begin{aligned} E[(X+Y)^2] &= E[X^2 + Y^2 + 2XY] = E[X^2] + E[Y^2] + 2E[XY] \\ &= E[X^2] + E[Y^2] + 2E[X]E[Y] \end{aligned}$$

Variance of sum of independent random variables

Let X and Y be two independent random variables. What is $\mathrm{var}(X+Y)$?

- Remember! $\mathrm{var}(X+Y) = E[(X+Y)^2] - (E[X+Y])^2$
- $$\begin{aligned} E[(X+Y)^2] &= E[X^2+Y^2+2XY] = E[X^2]+E[Y^2]+2E[XY] \\ &= E[X^2]+E[Y^2]+2E[X]E[Y] \\ E[X+Y]^2 &= (E[X]+E[Y])^2 = E[X]^2+E[Y]^2+2E[X]E[Y] \\ \mathrm{var}(X+Y) &= E[(X+Y)^2] - (E[X+Y])^2 \end{aligned}$$

Variance of sum of independent random variables

Let X and Y be two independent random variables. What is $\mathrm{var}(X+Y)$?

- Remember! $\mathrm{var}(X+Y) = E[(X+Y)^2] - (E[X+Y])^2$
- $$\begin{aligned} E[(X+Y)^2] &= E[X^2 + Y^2 + 2XY] = E[X^2] + E[Y^2] + 2E[XY] \\ &= E[X^2] + E[Y^2] + 2E[X]E[Y] \\ E[X+Y]^2 &= (E[X] + E[Y])^2 = E[X]^2 + E[Y]^2 + 2E[X]E[Y] \\ \mathrm{var}(X+Y) &= E[(X+Y)^2] - (E[X+Y])^2 \\ &= \underbrace{E[X^2] - E[X]^2}_{\mathrm{var}(X)} + \underbrace{E[Y^2] - E[Y]^2}_{\mathrm{var}(Y)} = \mathrm{var}(X) + \mathrm{var}(Y) \end{aligned}$$

Variance of sum of independent random variables

Let X and Y be two independent random variables. What is $\mathrm{var}(X+Y)$?

- Remember! $\mathrm{var}(X+Y) = E[(X+Y)^2] - (E[X+Y])^2$
- $$\begin{aligned} E[(X+Y)^2] &= E[X^2 + Y^2 + 2XY] = E[X^2] + E[Y^2] + 2E[XY] \\ &= E[X^2] + E[Y^2] + 2E[X]E[Y] \\ E[X+Y]^2 &= (E[X] + E[Y])^2 = E[X]^2 + E[Y]^2 + 2E[X]E[Y] \\ \mathrm{var}(X+Y) &= E[(X+Y)^2] - (E[X+Y])^2 \\ &= \underbrace{E[X^2] - E[X]^2}_{\mathrm{var}(X)} + \underbrace{E[Y^2] - E[Y]^2}_{\mathrm{var}(Y)} = \mathrm{var}(X) + \mathrm{var}(Y) \end{aligned}$$
- Variance of sum of independent random variables equals the sum of the variances!

Independence of several random variables

- Three random variables X, Y and Z are said to be independent if

$$P(X = x, Y = y, Z = z) = P(X = x)P(Y = y)P(Z = z) \qquad \text{For all } x, y, z$$

- If X, Y, Z are independent, then so are $f(X)$, $g(Y)$ and $h(Z)$.
- Also, any random variable $f(X, Y)$ and $g(Z)$ are independent.
- Are $f(X, Y)$ and $g(Y, Z)$ independent?

Independence of several random variables

- Three random variables X, Y and Z are said to be independent if

$$P(X = x, Y = y, Z = z) = P(X = x)P(Y = y)P(Z = z) \qquad \text{For all } x, y, z$$

- If X, Y, Z are independent, then so are $f(X)$, $g(Y)$ and $h(Z)$.
- Also, any random variable $f(X, Y)$ and $g(Z)$ are independent.
- Are $f(X, Y)$ and $g(Y, Z)$ independent?
  - Not necessarily, both have Y in common.

Independence of several random variables

- Three random variables X, Y and Z are said to be independent if

$$P(X = x, Y = y, Z = z) = P(X = x)P(Y = y)P(Z = z) \qquad \text{For all } x, y, z$$

- If X, Y, Z are independent, then so are $f(X)$, $g(Y)$ and $h(Z)$.
- Also, any random variable $f(X, Y)$ and $g(Z)$ are independent.
- Are $f(X, Y)$ and $g(Y, Z)$ independent?
  - Not necessarily, both have Y in common.

- For n independent random variables, X_1, $X_2, \ldots, X_n$, we also have:

$$\mathrm{var}(X_1 + X_2 + X_3 + \cdots + X_n) = \mathrm{var}(X_1) + \mathrm{var}(X_2) + \cdots + \mathrm{var}(X_n)$$

Variance of a Binomial

Consider n independent Bernoulli variables $X_1, X_2, \ldots, X_n$, each with probability p of having value "1". The sum $Y = \sum_i X_i$ is a $Binomial(n, p)$ random variable.

- We saw last time that $E[Y] = \sum_i E[X_i] = np$. What about the variance?
- Recall that $\mathrm{var}(X_i) = p(1-p)$ for $i \in \{1, 2, \ldots, n\}$.
- $\mathrm{var}(Y) = \mathrm{var}(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^{n} \mathrm{var}(X_i) = np(1-p)$.

Conditional Independence

- Very similar to conditional independence of events!
- X and Y are conditionally independent, given a positive probability event A if

$$P(X = x, Y = y|A) = P(X = x|A)P(Y = y|A) \qquad \text{For all } x \text{ and } y$$

- Same as saying $P(X = x|Y = y, A) = P(X = x|A)$, i.e.
- Once you know that A has occurred, knowing $\{Y = y\}$ has occurred does not give you any more information!
- Like we learned before, conditional independence does not imply unconditional independence.

Example-conditionally independent but not marginally

- I separately phone two students (Alice and Bob) and tell them the midterm grade.
- To each, I report the same grade, $G \in \{A+, A..., C\}$.
- The signal is bad and, Alice and Bob each independently make an educated guess of what I said.
- Let the grades guessed by Alice and Bob be X and Y.
- Are X and Y marginally independent?

Example-conditionally independent but not marginally

- I separately phone two students (Alice and Bob) and tell them the midterm grade.
- To each, I report the same grade, $G \in \{A+, A..., C\}$.
- The signal is bad and, Alice and Bob each independently make an educated guess of what I said.
- Let the grades guessed by Alice and Bob be X and Y.
- Are X and Y marginally independent?
  - NO. you would think, $P(X = A|Y = A) > P(X = A)$.

Example-conditionally independent but not marginally

- I separately phone two students (Alice and Bob) and tell them the midterm grade.
- To each, I report the same grade, $G \in \{A+, A..., C\}$.
- The signal is bad and, Alice and Bob each independently make an educated guess of what I said.
- Let the grades guessed by Alice and Bob be X and Y.
- Are X and Y marginally independent?
  - NO. you would think, $P(X = A|Y = A) > P(X = A)$.
- What if I tell you that $G = A-$?

Example-conditionally independent but not marginally

- I separately phone two students (Alice and Bob) and tell them the midterm grade.
- To each, I report the same grade, $G \in \{A+, A..., C\}$.
- The signal is bad and, Alice and Bob each independently make an educated guess of what I said.
- Let the grades guessed by Alice and Bob be X and Y.
- Are X and Y marginally independent?
  - NO. you would think, $P(X = A|Y = A) > P(X = A)$.
- What if I tell you that $G = A-$?
  - Are X and Y conditionally independent given $\{G = A-\}$.

Example-conditionally independent but not marginally

- I separately phone two students (Alice and Bob) and tell them the midterm grade.
- To each, I report the same grade, $G \in \{A+, A..., C\}$.
- The signal is bad and, Alice and Bob each independently make an educated guess of what I said.
- Let the grades guessed by Alice and Bob be X and Y.
- Are X and Y marginally independent?
  - NO. you would think, $P(X = A|Y = A) > P(X = A)$.
- What if I tell you that $G = A-$?
  - Are X and Y conditionally independent given $\{G = A-\}$.
  - YES! Because if we know the grade I actually said, the two variables are no longer dependent.

Example-marginally independent but not conditionally

- I toss two dice independently and X and Y are the readings on them.
- Are X and Y independent?
- Now I tell you that $X + Y = 12$. Are they still independent?