

Course support and deadlines

Ed Discussion: questions, announcements, and private messages. **No Canvas Inbox.**

[Open Ed from Canvas](#) All times Pacific

[Full contact details and meeting links](#)

Instructor / TAs	Office hours	Optional, ungraded discussions
Antonio de Leon GSC (in person)	Mon 3–4 p.m. (extra) Fri 3–4 p.m.	—
Jeannine Romero	Tue / Thu 10–11 a.m.	01A: Fri 9:20–10:25 a.m. R Carson Acad 240 (in person)
Ehsan Mirafzali	Mon 3–5 p.m.	01B: Fri 10:40–11:45 a.m. (Zoom)
Seokjun Choi	Wed 3–4 p.m. 5:05–6:05 p.m.	01C: Wed 4–5:05 p.m. (Zoom)

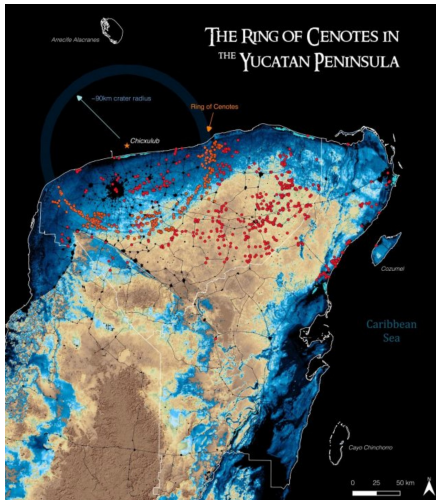
Upcoming deadlines

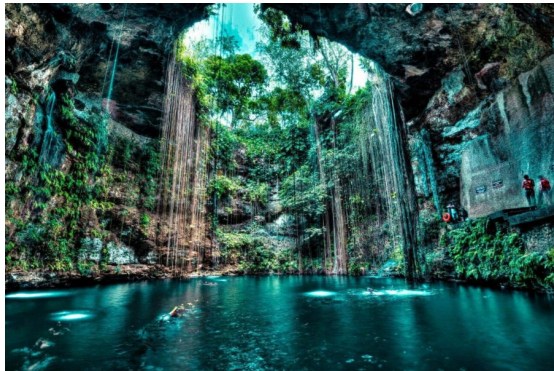
HW and quizzes: **10 p.m. Pacific**

Quiz 4	Sun, Oct 25	Canvas
HW 3	Tue, Oct 27	Gradescope

GSC: Graduate Student Commons. TA office-hour times [link to Zoom](#).

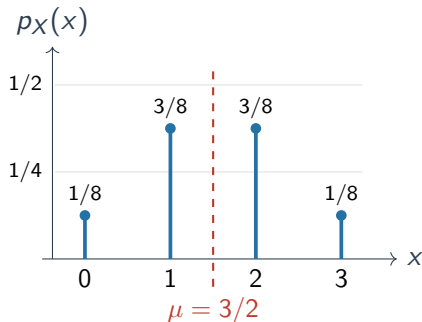
HW: 24-hour grace period. Quizzes: no grace period. Approved accommodations apply.





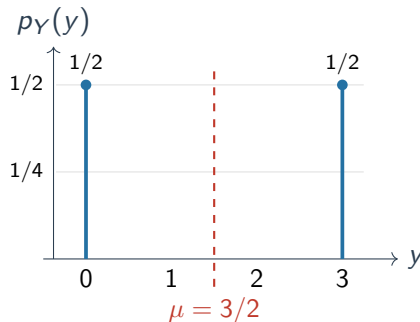
Three independent fair tosses

X counts heads: $X \sim \text{Bin}(3, 1/2)$.



A hypothetical all-or-none count

$Y = 0$ or 3 , each with probability $1/2$.



Both means are $3/2$. Which law puts more probability far from its mean?

These are two separate count models; no joint relationship is assumed.

Expectation of a function

Finite support S_X and a real-valued function g

$$\mathbb{E}[g(X)] = \sum_{x \in S_X} g(x)p_X(x).$$

Derivation. Put $Z = g(X)$. Group all inputs x giving the same output z , then substitute into $\mathbb{E}[Z] = \sum_z zP(Z = z)$.

Evaluate g at each possible value, then average with the original weights.

We prove the finite case. For countable support, require $\sum_{x \in S_X} |g(x)|p_X(x) < \infty$.

First and second moments

For $X \sim \text{Bin}(3, 1/2)$, with $S_X = \{0, 1, 2, 3\}$, complete the two weighted sums.

x	0	1	2	3	Sum
$p_X(x)$	1/8	3/8	3/8	1/8	1
$x p_X(x)$					
$x^2 p_X(x)$					

$$\mathbb{E}[X^2] = \sum_{x \in S_X} x^2 p_X(x), \quad (\mathbb{E}[X])^2 = \left(\sum_{x \in S_X} x p_X(x) \right)^2.$$

The second moment averages the squared values.

Variance and standard deviation

Assume $\mathbb{E}[X^2] < \infty$ and write $\mu = \mathbb{E}[X]$.

Expected squared deviation

$$\sigma^2 = \text{Var}(X) = \mathbb{E}[(X - \mu)^2], \quad \text{SD}(X) = \sigma = \sqrt{\text{Var}(X)}.$$

Why square? Signed deviations cancel: $\mathbb{E}[X - \mu] = 0$.

Finite-PMF argument. Explain nonnegativity and the zero-variance case:

$$\text{Var}(X) = \sum_{x \in S_X} (x - \mu)^2 p_X(x).$$

If X is measured in milliseconds, variance has units ms^2 and SD has units ms .

Computing variance from two moments

For $\mathbb{E}[X^2] < \infty$

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

Derive it from the definition. Keep $\mu = \mathbb{E}[X]$ fixed as a number.

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

Expand the square, use linearity, and substitute $\mathbb{E}[X] = \mu$.

Linearity applies to the expectation in the definition. Independence is not used.

Variance of the coin count

For $X \sim \text{Bin}(3, 1/2)$, we found $\mathbb{E}[X] = 3/2$ and $\mathbb{E}[X^2] = 3$.

Using centered values

$$\text{Var}(X) = \sum_{x=0}^3 (x - 3/2)^2 p_X(x).$$

Using two moments

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

The methods agree. Equal means do not imply equal variances.

Shifts and changes of scale

Let $Z = aX + b$, with real constants a, b and $\mathbb{E}[X^2] < \infty$.

Affine change

$$\text{Var}(aX + b) = a^2 \text{Var}(X), \quad \text{SD}(aX + b) = |a| \text{SD}(X).$$

Center first. With $\mu = \mathbb{E}[X]$, linearity gives $\mathbb{E}[Z] = a\mu + b$.

A shift leaves centered deviations unchanged; scaling multiplies their squares by a^2 .

If $a = 0$, then $Z = b$ is constant and its variance is zero.

A delay measured in milliseconds

Hypothetical model. Each of three independent subtasks needs an extra operation with probability $1/2$. Let X count these extra operations.

$$L = 100 + 20X \text{ ms}, \quad X \sim \text{Bin}(3, 1/2).$$

Use $\mathbb{E}[X] = 3/2$ and $\text{Var}(X) = 3/4$.

Quantity	Calculation	Units
$\mathbb{E}[L]$		ms
$\text{Var}(L)$		ms ²
$\text{SD}(L)$		ms

Check the units as well as the number.

Bernoulli variance

Let $I \sim \text{Bern}(p)$, where $0 \leq p \leq 1$.

$$P(I = 1) = p, \quad P(I = 0) = 1 - p, \quad \mathbb{E}[I] = p.$$

Use the support. Since $I \in \{0, 1\}$, we have $I^2 = I$.

One trial

$$\mathbb{E}[I] = p, \quad \text{Var}(I) = p(1 - p), \quad \text{SD}(I) = \sqrt{p(1 - p)}.$$

Compute and rescale a variance

w	0	1	3
$p_W(w)$	1/2	1/4	1/4

Find $\mathbb{E}[W]$, $\mathbb{E}[W^2]$, $\text{Var}(W)$, and $\text{Var}(5 - 2W)$.

For SD, the scale factor is $|-2| = 2$.

Variance summary

Meaning	$\text{Var}(X) = \mathbb{E}[(X - \mu)^2]$
Computation	$\text{Var}(X) = \mathbb{E}[X^2] - \mu^2$
Affine change	$\text{Var}(aX + b) = a^2 \text{Var}(X)$
Standard deviation	$\text{SD}(aX + b) = a \text{SD}(X)$

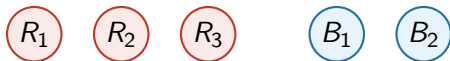
Here $\mu = \mathbb{E}[X]$ and $\mathbb{E}[X^2] < \infty$.

Variance is not linear

For $I \sim \text{Bern}(1/2)$,

$$\text{Var}(2I) = 1 \quad \text{but} \quad 2 \text{Var}(I) = \frac{1}{2}.$$

Draw two balls uniformly **without replacement** from three red and two blue balls.



$$X = I_{\{\text{first draw is red}\}}, \quad Y = I_{\{\text{second draw is red}\}}.$$

Each indicator equals 1 for red and 0 for blue. Here X and Y describe the two draws.

Physical outcome	Color pair	(X, Y)
(R_2, B_1)	red, blue	$(1, 0)$
(B_2, R_3)	blue, red	$(0, 1)$

The two variables describe the same outcome. Independence has not been assumed.

Joint probability mass functions

Discrete X and Y

$$p_{X,Y}(x,y) = P(X = x, Y = y).$$

The events $\{X = x, Y = y\}$ form disjoint cells covering the sample space.

$$p_{X,Y}(x,y) \geq 0, \quad \sum_x \sum_y p_{X,Y}(x,y) = 1.$$

Justify normalization by adding the probabilities of the disjoint cells.

The joint support is the set of pairs with positive mass; put zero outside it.

All worked tables are finite. These nonnegative-sum formulas also hold for countable supports.

Two draws without replacement

There are $5 \cdot 4 = 20$ equally likely ordered pairs of distinct physical balls.

Count the pairs in each cell

$X \backslash Y$	0 (blue)	1 (red)
0 (blue)	$2 \cdot 1$	$2 \cdot 3$
1 (red)	$3 \cdot 2$	$3 \cdot 2$

Normalize by the total count

$X \backslash Y$	0	1
0		
1		

The four counts are 2, 6, 6, 6.

Equal probability applies to physical-ball pairs, not automatically to color pairs.

Events in a joint table

Add the masses of the cells in the event:

$$P((X, Y) \in A) = \sum_{(x,y) \in A} p_{X,Y}(x,y).$$

The colors agree: $X = Y$

$X \backslash Y$	0	1
0	1/10	3/10
1	3/10	3/10

At least one red: $X + Y \geq 1$

$X \backslash Y$	0	1
0	1/10	3/10
1	3/10	3/10

Translate the event into cells before adding their masses.

Deriving a marginal PMF

Fix x . The possible values of Y split the event $\{X = x\}$:

$$\{X = x\} = \bigcup_y \{X = x, Y = y\} \quad (\text{disjoint union}).$$

Marginal distributions

$$p_X(x) = \sum_y p_{X,Y}(x, y), \quad p_Y(y) = \sum_x p_{X,Y}(x, y).$$

Keep the variable you want; sum over the variable you remove.

Row totals and column totals

Complete the margins of the two-draw table.

$X \backslash Y$	0	1	$p_X(x)$
0	1/10	3/10	
1	3/10	3/10	
$p_Y(y)$			1

Check: each marginal has nonnegative entries summing to one.

Equal marginal distributions do not establish independence.

The total number of red balls

Let $N = X + Y$. Group cells by their total, then recover $\mathbb{E}[N]$ in two ways.

n	Pairs (x, y)	$p_N(n)$
0	(0, 0)	1/10
1	(0, 1), (1, 0)	
2	(1, 1)	3/10

Linearity gives the same answer even though these draws are dependent.

The same marginals, two joint laws

X, Y indicate red on two draws from three red and two blue balls.

Without replacement

$X \backslash Y$	0	1	$p_X(x)$
0	1/10	3/10	2/5
1	3/10	3/10	3/5
$p_Y(y)$	2/5	3/5	1

Independent draws with replacement

$X \backslash Y$	0	1	$p_X(x)$
0	4/25	6/25	2/5
1	6/25	9/25	3/5
$p_Y(y)$	2/5	3/5	1

Compare the probability of two reds.

Marginals alone do not determine the joint law.

Marginals and an event probability

$X \setminus Y$	0	1	2	$p_X(x)$
0	.10	.25	.35	
1	.15	.05	.10	
$p_Y(y)$				1

Find both marginals and $P(Y > X)$. Mark the selected cells.

Joint PMFs and marginals

- 1 Define the experiment and both variables.
- 2 Label the possible values and enter zero for impossible pairs.
- 3 Check every entry is nonnegative and the grand total is one.
- 4 Sum rows for p_X and columns for p_Y .
- 5 Check each marginal sums to one.

Three different objects

An outcome belongs to the sample space.

A joint cell records a pair of values (x, y) .

A marginal law records one variable's values and probabilities.

Recall the two draws **without replacement**.

$X \backslash Y$	0	1	$p_X(x)$
0	1/10	3/10	2/5
1	3/10	3/10	3/5
$p_Y(y)$	2/5	3/5	1

Describe each event and find its probability.

$$P(X = 1, Y = 1), \quad P(X = 1).$$

Now suppose the first draw was red. What is the probability the second is red?

The observed first draw restricts us to the row $X = 1$.

Deriving a conditional PMF

Fix x with $p_X(x) > 0$

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

This is event conditioning with $A = \{Y = y\}$ and $B = \{X = x\}$.

Show that it is a PMF. Its entries are nonnegative. Check their sum.

Select a row, then divide every entry by that row's probability.

If $p_X(x) = 0$, this elementary ratio does not define a conditional PMF.

Condition on the first draw

Select the row $X = 1$

$X \backslash Y$	0	1	$p_X(x)$
0	1/10	3/10	2/5
1	3/10	3/10	3/5
$p_Y(y)$	2/5	3/5	1

Normalize by $p_X(1) = 3/5$

y	0	1
$p_{Y X}(y 1)$		

Compare with the marginal: $p_Y(1) = 3/5$.

Here observing $X = 1$ changes the law of Y . Under independence it would not.

Reversing the conditioning

Use the asymmetric joint table from Wednesday. Select the conditioning event first.

Fix the row $X = 1$

y	0	1	2
$p_{X,Y}(1, y)$	.15	.05	.10

The row total is $p_X(1) = .30$.

$$P(Y = 2 \mid X = 1) = \frac{.10}{.30}.$$

Fix the column $Y = 2$

x	$p_{X,Y}(x, 2)$
0	.35
1	.10

The column total is $p_Y(2) = .45$.

$$P(X = 1 \mid Y = 2) = \frac{.10}{.45}.$$

The numerator is the same cell; the conditioning event determines the denominator.

Recovering joint and marginal probabilities

For $p_X(x) > 0$, multiply the conditional-PMF definition by its denominator:

$$p_{X,Y}(x, y) = p_X(x)p_{Y|X}(y | x).$$

Then sum over x to recover the marginal.

$$p_Y(y) = \sum_{x \in S_X} p_X(x)p_{Y|X}(y | x), \quad S_X = \{x : p_X(x) > 0\}.$$

Without replacement: after blue, 3 of 4 remaining balls are red; after red, 2 of 4.

Average conditional probabilities with their event probabilities as weights.

Rows with $p_X(x) = 0$ contain only zero joint masses and contribute nothing.

Independence and factorization

Independence means $P(X \in A, Y \in B) = P(X \in A) P(Y \in B)$ for all value sets A, B .

Equivalent criterion for discrete variables

$$p_{X,Y}(x,y) = p_X(x)p_Y(y) \quad \text{for every } (x,y) \in S_X \times S_Y.$$

Here S_X, S_Y are the marginal supports. Test zero joint cells too.

Finite proof. Use singleton events for one direction. For the converse, sum over $A \times B$.

Conditional laws under independence

Equivalent conditional criterion

X and Y are independent exactly when, for every x with $p_X(x) > 0$,

$$p_{Y|X}(y | x) = p_Y(y) \quad \text{for every } y.$$

Derive both directions by division or multiplication by $p_X(x)$.

One unequal cell refutes independence. One agreeing cell does not prove it.

Independence requires positive joint mass at every pair in $S_X \times S_Y$. This condition alone is insufficient.

Replacement changes the joint law

Without replacement

$X \setminus Y$	0	1	$p_X(x)$
0	1/10	3/10	2/5
1	3/10	3/10	3/5
$p_Y(y)$	2/5	3/5	1

Independent draws with replacement

$X \setminus Y$	0	1	$p_X(x)$
0	4/25	6/25	2/5
1	6/25	9/25	3/5
$p_Y(y)$	2/5	3/5	1

Both tables have positive mass in every cell. Only the replacement table factors.

Does this table show independence?

$X \backslash Y$	0	1	2	$p_X(x)$
0	.12	.18	.30	
1	.08	.12	.20	
$p_Y(y)$				1

Compute the marginals and test every cell. Then find $P(Y \geq 1 \mid X = 0)$.

Joint, marginal, and conditional PMFs

Object	Formula	Operation
Joint	$p_{X,Y}(x, y) = P(X = x, Y = y)$	Select one cell
Marginal	$p_X(x) = \sum_y p_{X,Y}(x, y)$	Sum a row
Conditional	$p_{Y X}(y x) = \frac{p_{X,Y}(x, y)}{p_X(x)}$	Normalize a row

The conditional formula requires $p_X(x) > 0$.

Independence

The joint equals the product of marginals for every pair of marginal values. Equivalently, every positive-probability conditional row reproduces p_Y .

Optional: the second moment of a binomial count

Let $I_1, \dots, I_n$ be independent $\text{Bern}(p)$ variables, and $N = \sum_i I_i$.

$$N^2 = \sum_{i=1}^n I_i^2 + 2 \sum_{i < j} I_i I_j.$$

Use $I_i^2 = I_i$. For $i \neq j$, $I_i I_j$ is the indicator of both successes.

Independence is used for the pair probability, not for linearity of expectation.

Optional: binomial variance

For $N \sim \text{Bin}(n, p)$, the indicator calculation gives

$$\mathbb{E}[N] = np, \quad \mathbb{E}[N^2] = np + n(n-1)p^2.$$

Apply the variance identity and simplify.

Binomial moments

$$\mathbb{E}[N] = np, \quad \text{Var}(N) = np(1-p).$$

Pairwise independence suffices for this variance calculation. The binomial trial model assumes mutual independence. General variance-of-sums rules are taught later.

Optional: the mean minimizes squared error

Assume $\mathbb{E}[X^2] < \infty$ and write $\mu = \mathbb{E}[X]$. For any real constant c ,

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

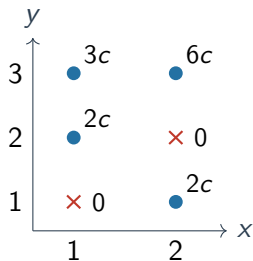
Derive it: write $X - c = (X - \mu) + (\mu - c)$, expand, and take expectation.

Variance is the smallest expected squared error from using a constant prediction.

Optional: support matters for independence

Let $S = \{(x, y) : x \in \{1, 2\}, y \in \{1, 2, 3\}, x \neq y\}$.

$$p_{X,Y}(x, y) = cxyI_{\{(x,y) \in S\}}.$$



Normalize, find marginals, test (1, 1).

The factor xy alone does not establish independence. Check the support too.

Reading selections

DeGroot and Schervish, *Probability and Statistics*, 4th ed.

- Variance: §4.1, p. 213 (discrete function-expectation proof); §4.3, pp. 226–229 (basic properties).
- Joint laws: §3.4, pp. 118–120; §3.5, pp. 131–132 (discrete marginals).
- Independence/conditioning: §3.5, pp. 134–136; §3.6, pp. 142–143 (discrete parts).

Harchol-Balter, *Introduction to Probability for Computing*, 2024 version.

- Moments/variance: p. 83 (definition); pp. 85–87 (variance and SD); p. 88 (Lemma 5.7 only).
- Joint PMFs/independence: §3.3, pp. 50–51; p. 53 (total-probability statement).
- Conditional PMFs: §4.3, pp. 67–68, stopping before conditional expectation.

Printed page numbers. Selections only; skip continuous arguments and later topics. Lectures are self-contained. Additional reference: Sheng Jiang, STAT 131 Notes (July 18, 2024), pp. 34, 38–43, 54–55, selected discrete statements.