LECTURE 14

Expected Value & Variance

Numerical functions of random samples and their properties

CSCI 3022 @ CU Boulder

Maribeth Oscamou

Announcements

- HW 6 Released Tonight
- Nb6 released Monday - presented next Tuesday in TA discussion

Some Goals of Data Science

- Understand the world better
- Help make the world better

For example

- Help expose injustice
- Help counter injustice

The skills that you have gained empower you to do this.

Discrete Random Variables

Common Discrete Random Variables

Bernoulli(p)

- Takes on value 1 with probability p, and 0 with probability 1 - p
- AKA the "indicator" random variable.

Binomial(n, p)

- Number of 1s in n independent Bernoulli(p) trials
- Probabilities given by the binomial formula

Poisson(u)

- Number of random events per unit of measurement

We'll use these in our class.
The rest are provided for your reference.

Uniform on a finite set of values

- Probability of each value is 1 / (size of set)
- For example, a standard die

Geometric(p)

The numbers in parentheses are the **parameters** of a random variable, which are constants. Parameters define a random variable's shape (i.e., distribution) and its values.

First Example

- U.S. Constitution grants equal protection under the law
- All defendants have the right to due process

We will study a U.S. Supreme Court case in the 1960s

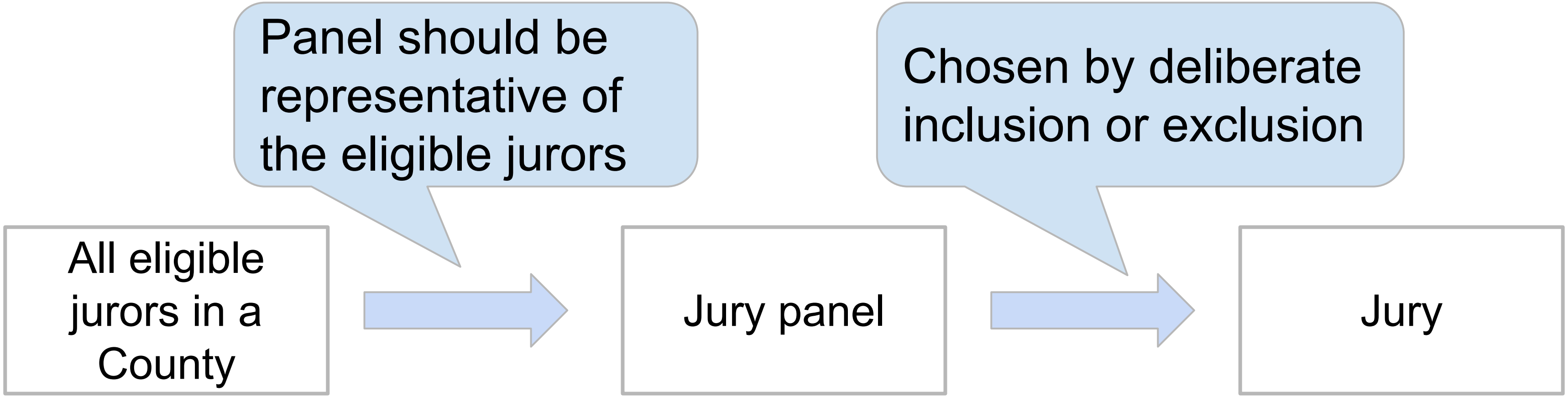
- A Black defendant was denied his Constitutional right to a fair jury
- The Court made incorrect and biased judgments about
    - the data in the case
    - the legal processes in the defendant's original trial
- We will discuss errors and racial bias in the Court's judgment

This case became the foundation of significant reform.

Jury Selection

US Constitution:

“right to a speedy and public trial, by an impartial jury”



Supreme Court Case

Robert Swain's Case

- Robert Swain, a Black man, was convicted in Talladega County, AL
- He appealed to the U.S. Supreme Court
- Main reason: Unfair jury selection in the County's trials

Panel should be representative of the eligible jurors

Chosen by deliberate inclusion or exclusion

Eligible jurors:
26% Black

→

Jury panel:
8 out of 100
Black

→

Jury:
0 Black

Supreme Court Ruling, 1965

- The Court denied Robert Swain's appeal.

Panel should be representative of the eligible jurors

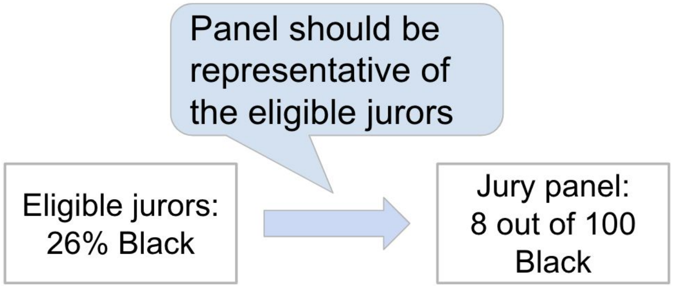
Eligible jurors: 26% Black

Jury panel: 8 out of 100 Black

About discrepancies like this, the Court wrote:

"the overall percentage disparity has been small"

Discussion Question



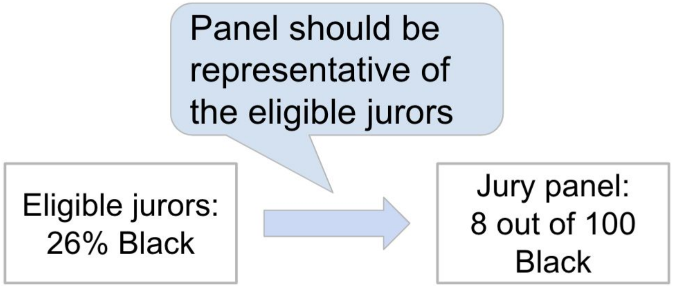
- Court's view: 8/100 is less than 26%, but not different enough to show Black panelists were systematically excluded

- Question: Would 8/100 be a realistic outcome if the jury panel selection process were truly unbiased?

- Let X be a random variable that describes the number of black jurors on a randomly selected panel of 100 eligible jurors in Talladega County, AL.

  What is the probability distribution of X?

Discussion Question



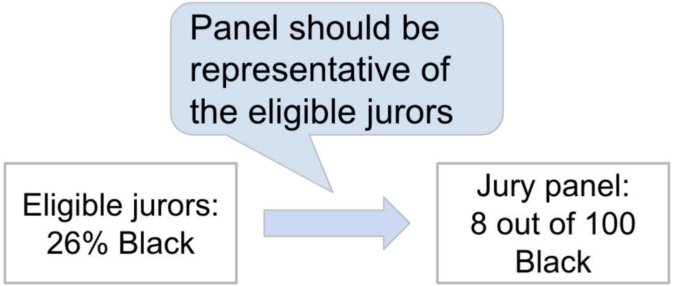
- Let X be a random variable that describes the number of black jurors on a randomly selected panel of 100 eligible jurors in Talladega County, AL.

  What is the probability distribution of X?

  What is $P(X=8)$?

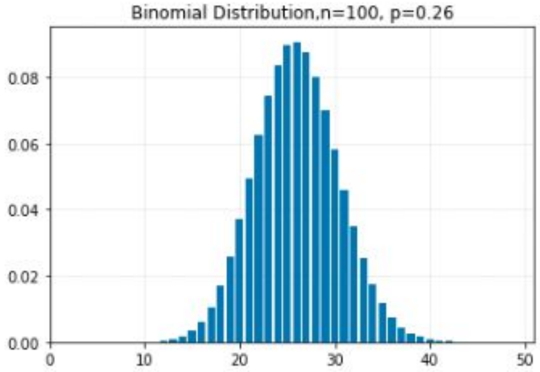
  What is $P(X<=8)$?

Discussion Question



- Let X be a random variable that describes the number of black jurors on a randomly selected panel of 100 eligible jurors in Talladega County, AL.

  What is the probability distribution of X?



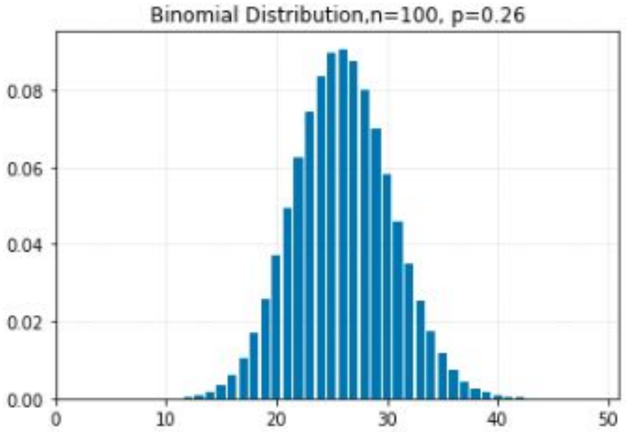
  What is P(X=8)?

  What is P(X<=8)?

Statistical Bias

- Evidence provided by Robert Swain:

  “only 10 to 15% of … jury panels drawn from the jury box since 1953 have been [Black], there having been only one case in which the percentage was as high as 23%”
- Percent of Black panelists was always lower than expected under random sampling
- *Bias*: when errors are systematically in one direction



Simulating Distributions

- Any random quantity has a probability distribution:
  - All possible values it can take
  - The probability it takes each value
    - Often challenging to calculate analytically (the math may not be possible…)

- After repeated draws, it has an empirical distribution:
  - All observed values it took
  - The proportion of times it took each value

- After many independent draws, the **empirical distribution** looks more and more like the ***probability distribution***

Lec14 NB Demo

Expectation and Variance

Should you gamble?

A Las Vegas roulette board contains 38 numbers $\{0, 00, 1, 2, \ldots, 36\}$. Of the non-zero numbers, 18 are red and 18 are black. You can place bets on various number/color combinations and each type of bet pays-out at a different rate. For example:

- If you bet $\$1$ on red (or black) and win, then you win $\$1$ (i.e. you get your original dollar back, plus another dollar).
- If you bet $\$1$ any particular number and win, then you win $\$35$ (i.e you get your original dollar back, plus $\$35$).
- If you bet $\$1$ on the first dozen (or second dozen, or third dozen) nonzero numbers and win, then you win $\$2$ (i.e. you get your original dollar back, plus another $\$2$).

You decide to bet $1 on any particular number.

You repeat this (for any particular number) 100 times and write down your net winnings for each trial.

What number below is the closest to the number you'd expect to get as the average net winnings in your 100 trials?

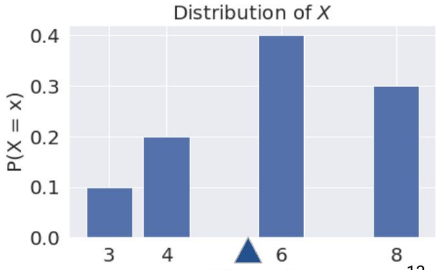
Definition of Expectation: Discrete Random Variables

The **expectation (or expected value)** of a random variable X is the **weighted average** of the values of X, where the weights are the probabilities of the values.

$$\mathbb{E}[X] = \sum_{\substack{\text{all possible} \\ x}} x P(X = x)$$

Expectation is a number, not a random variable!

- It is the center of gravity or balance point of the probability histogram.
- It is analogous to the mean: If we had exactly 100 students and their proportions observed exactly mirrored the probabilities given by the probability distribution of X, then the sample mean of our data would be identical to the expected value of X.



x	P(X = x)
3	0.1
4	0.2
6	0.4
8	0.3

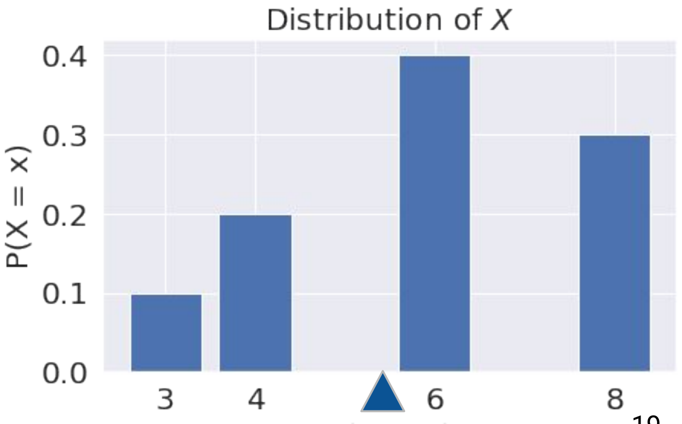
Example

$$\mathbb{E}[X] = \sum_x xP(X = x)$$

Consider the random variable X we defined last time:

x	P(X = x)
3	0.1
4	0.2
6	0.4
8	0.3

$$\mathbb{E}[X] = \sum_x x \cdot P(X = x)$$
$$= 3 \cdot 0.1 + 4 \cdot 0.2 + 6 \cdot 0.4 + 8 \cdot 0.3$$
$$= 0.3 + 0.8 + 2.4 + 2.4$$
$$= 5.9$$



Note, $E[X] = 5.9$ is not a possible value of X!
It is an average.

The expectation of X does not need to be a value of X.

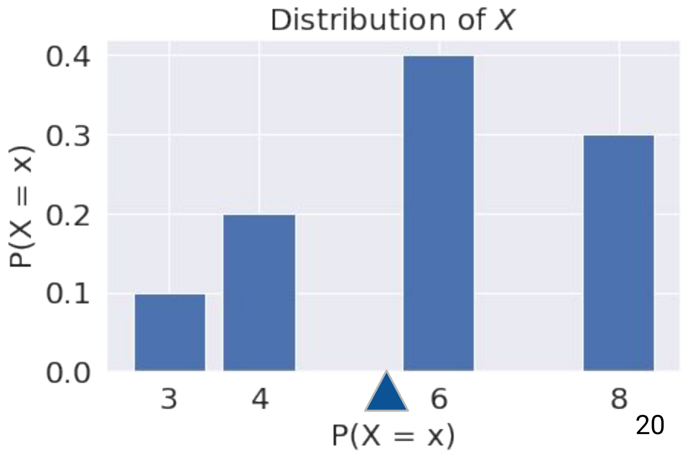
Example

$$\mathbb{E}[X] = \sum_x xP(X = x)$$

Consider the random variable X we defined last time:

$$\mathbb{E}[X] = \sum_x x \cdot P(X = x)$$
$$= 3 \cdot 0.1 + 4 \cdot 0.2 + 6 \cdot 0.4 + 8 \cdot 0.3$$
$$= 0.3 + 0.8 + 2.4 + 2.4$$
$$= 5.9$$

x	P(X = x)
3	0.1
4	0.2
6	0.4
8	0.3

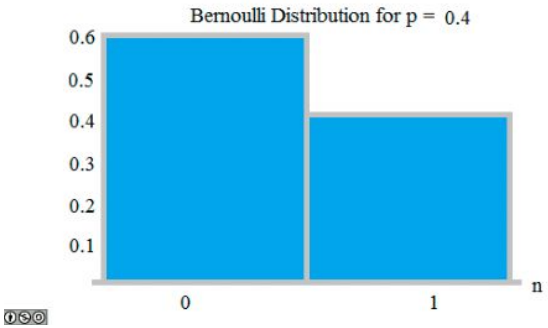


If we had exactly 100 students and their proportions observed exactly mirrored the probabilities in given in this example, the sample mean of our data would be identical to the expected value of X.

Expected Value of Bernoulli Distribution

A discrete random variable X has a **Bernoulli distribution,** denoted $X \sim Ber(p)$, where $0 \leq p \leq 1$, if its probability distribution is given by

$$P(X=1) = p$$
$$P(X=0) = 1-p$$



Examples of Bernoulli Random Variables:

- Outcome from flipping a coin
- Getting an answer correct when guessing on a multiple choice question
- Buying a winning lottery ticket
- Asking someone if they're left-handed

Let X~Ber(p)

What is E[X]?

Should you gamble?

A Las Vegas roulette board contains 38 numbers $\{0, 00, 1, 2, \ldots, 36\}$. Of the non-zero numbers, 18 are red and 18 are black. You can place bets on various number/color combinations and each type of bet pays-out at a different rate. For example:

- If you bet $\$1$ on red (or black) and win, then you win $\$1$ (i.e. you get your original dollar back, plus another dollar).
- If you bet $\$1$ any particular number and win, then you win $\$35$ (i.e you get your original dollar back, plus $\$35$).
- If you bet $\$1$ on the first dozen (or second dozen, or third dozen) nonzero numbers and win, then you win $\$2$ (i.e. you get your original dollar back, plus another $\$2$).

Let X be your winnings if you bet $1 on any particular number and play once.

a). What is the probability distribution of X?

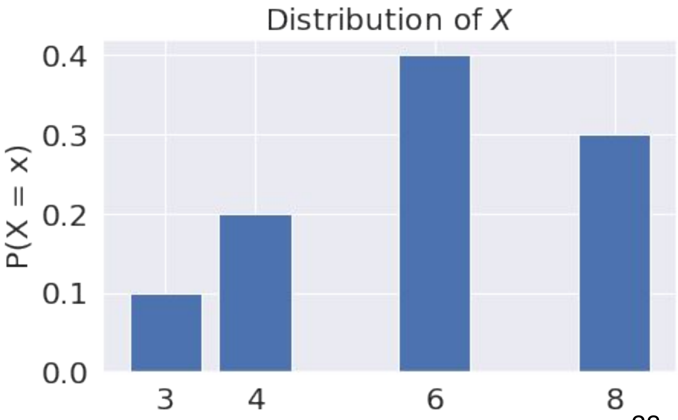
b). What is E[X]?

Expected Value of a Function of X:

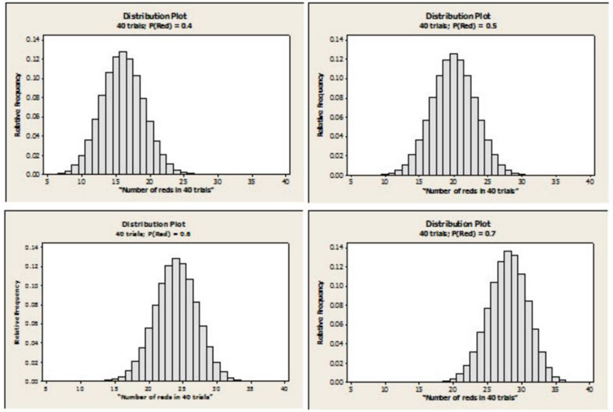
Consider the random variable X we defined earlier.

Calculate

x	P(X = x)
3	0.1
4	0.2
6	0.4
8	0.3



Expected Value of Binomial and Poisson Distributions

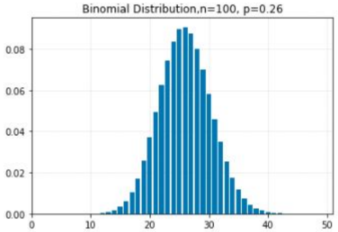


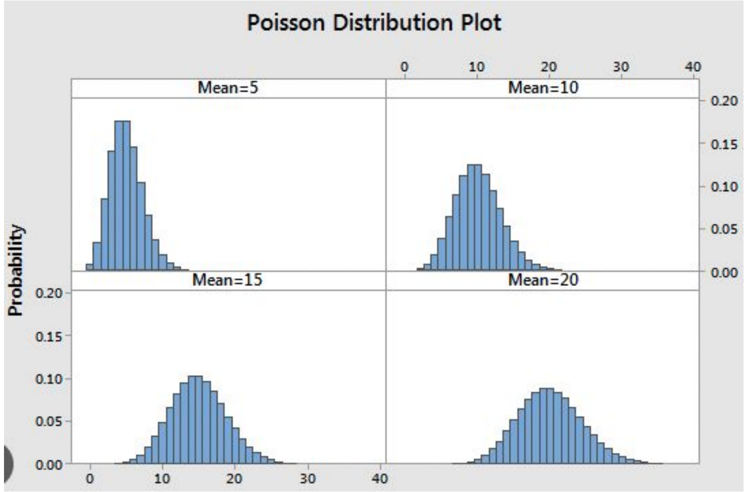
Binomial Distribution

A discrete random variable X has a **Binomial Distribution,** denoted $X \sim Bin(n, p)$, with $n = 1, 2, \ldots$ and $0 \le p \le 1$, if its probability distribution is given by

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } 0 \le k \le n$$

p = probability of success for each of the n trials
k = number of successful trials out of n total





Poisson Distribution

A discrete random variable X has a **Poisson distribution** denoted $X \sim Pois(\mu)$ with parameter $\mu > 0$, if its distribution is given by

$$P(X = k) = \frac{\mu^k e^{-\mu}}{k!} \quad \text{for } k = 0, 1, 2, \ldots$$

(i.e. probability of k events per unit of measurement)

$\boldsymbol{\mu}$ is the expected number of events per unit of measurement (time period, length, space, volume,etc)

Proof

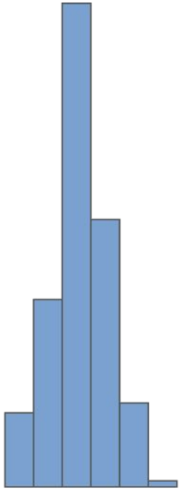
Defining Variability

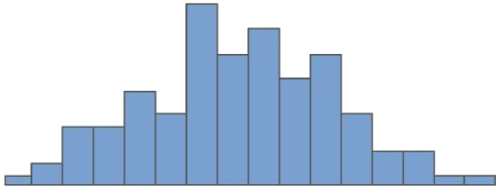
Plan A: "biggest value - smallest value"

- Doesn't tell us much about the shape of the distribution

Plan B:

- Measure variability around the mean
- Need to figure out a way to quantify this





Definition of Variance

Variance is the **expected squared deviation from the expectation** of X.

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

- The units of the variance are the square of the units of X.
- To get back to the right scale, use the **standard deviation** of X: $\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$

Definition of Variance

Variance is the **expected squared deviation from the expectation** of X.

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

- The units of the variance are the square of the units of X.
- To get back to the right scale, use the **standard deviation** of X: $\text{SD}(X) = \sqrt{\text{Var}(X)}$

Variance is a number, not a random variable!

- The main use of variance is to **quantify chance error**. How far away from the expectation could X be, just by chance?

There's a more convenient form of variance:

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

- Proof (involves expanding the square and properties of expectation/summations): link
- When computing variance by hand, often used instead of definition.

How Far from the Average?

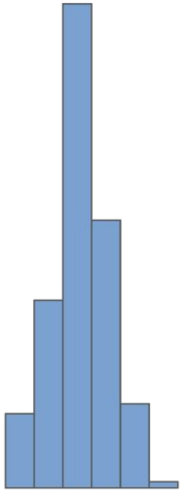
- Standard deviation (SD) measures roughly how far the data are from their average

- SD = Root Mean Square of Deviations from Average

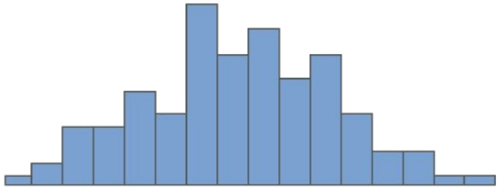
Steps: 5 ← 4 ← 3 ← 2 ← 1

(SD is known as the RMS of the deviations)

- SD has the same units as the data



Standard Deviation is <u>small</u>



Standard Deviation is <u>large</u>

No matter what the shape of the distribution,

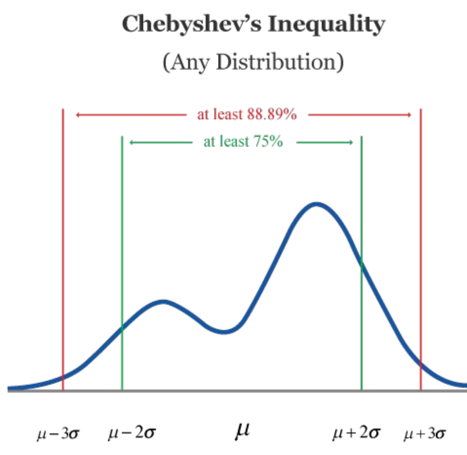
the bulk of the data are in the range "average ± a few SDs"

Chebyshev's Inequality

No matter what the shape of the distribution,

the proportion of values in the range "mean ± z SDs" is

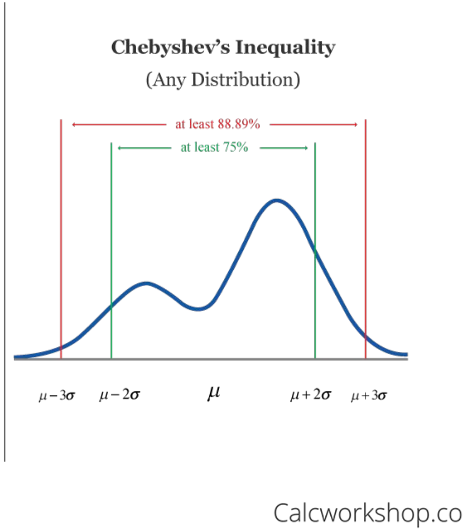
at least $1 - 1/z^2$



Calcworksh

Chebyshev's Bounds

Range	Proportion
average ± 2 SDs	at least 1 - 1/4 = 3/4 (75%)
average ± 3 SDs	at least 1 - 1/9 = 8/9 (88.88…%)
average ± 4 SDs	at least 1 - 1/16 = 15/16 (93.75%)
average ± 5 SDs	at least 1 - 1/25 = 24/25 (96%)



No matter what the distribution looks like!

Example

$$\mathbb{E}[X] = \sum_x x P(X = x)$$

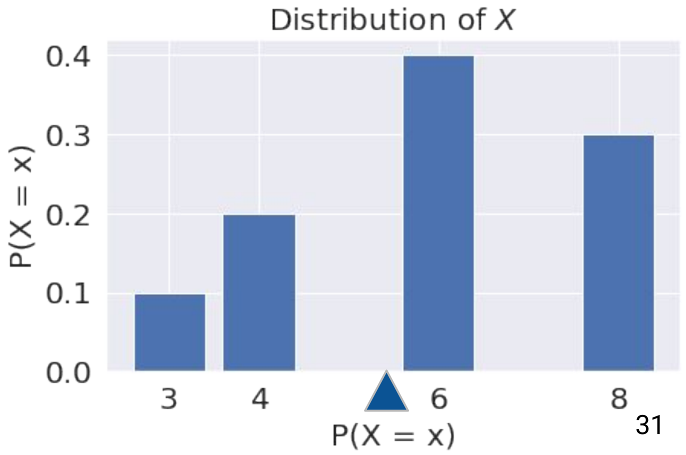
Consider the random variable X we defined earlier.

x	P(X = x)
3	0.1
4	0.2
6	0.4
8	0.3

What is the variance, $\text{Var}(X)$?

Approach 1: Definition

Approach 2: Property



Properties of Bernoulli Random Variables

$$\mathbb{E}[X] = \sum_x x P(X = x)$$

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \end{aligned}$$

Definitions

Let X be a **Bernoulli**(p) random variable.

- Takes on value 1 with probability p, and 0 with probability 1 - p.
- AKA the “indicator” random variable.

Expectation:

$$\mathbb{E}[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

We will get an average value of p across many, many samples

Variance:

$$\mathbb{E}[X^2] = 1^2 \cdot p + 0^{\cdot}(1 - p) = p$$

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= p - p^2 = p(1 - p) \end{aligned}$$

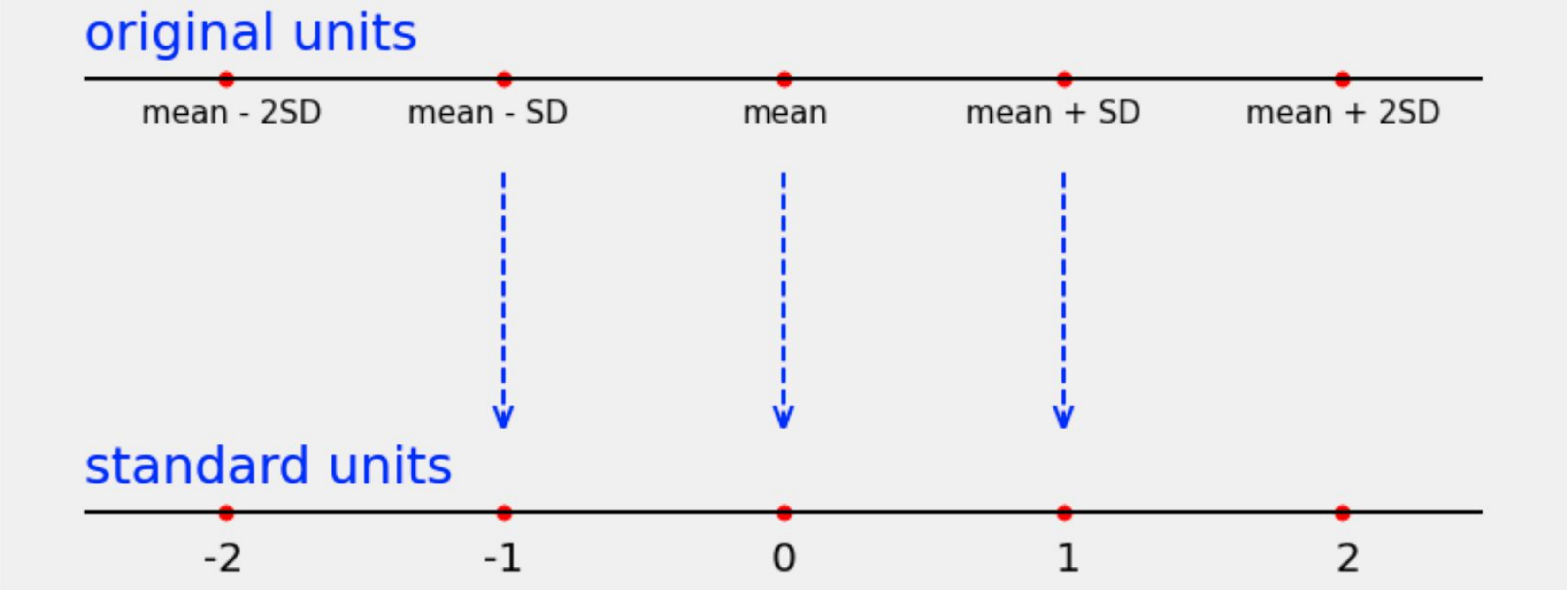
Lower Var: p = 0.1 or 0.9
Higher Var: p close to 0.5

More info: google(“plot x(1 - x)”)

Standard Units

Standard Units

standard units = (original value - mean) / SD



Standard Units

- **Measures:** How many SDs above average?

- *z* = (value - average)/SD
  - Negative z: value below average
  - Positive z: value above average
  - z = 0: value equal to average

(Demo)

- When values are in standard units: average = 0, SD = 1

Standardization of random variables

Jump back: link

X in **standard units** is the random variable $$X_{su} = \frac{X - \mathbb{E}(X)}{\mathbb{SD}(X)}.$$

X_{su} measures X on the scale **"number of SDs from expectation**."

- Since X_{su} is centered (has expectation 0):

$$\mathbb{E}(X_{su}) = 0, \quad \mathbb{SD}(X_{su}) = 1$$

$$\mathbb{E}(X_{su}^2) = \mathbb{V}ar(X_{su}) = 1$$

Discussion Question

Find whole numbers that are close to:

(a) the average age

(b) the SD of the ages

Age in Years	Age in Standard Units
27	-0.0392546
33	0.992496
28	0.132704
23	-0.727088
25	-0.383171
33	0.992496
23	-0.727088
25	-0.383171
30	0.476621
27	-0.0392546

... (1164 rows omitted)

Functions of Multiple Random Variables

A function of a random variable is also a random variable!

If you create multiple random variables based on your sample…

…then functions of those random variables are also random variables.

For instance, if $X_1, X_2, \ldots, X_n$ are random variables, then so are all of these:

$$X_n^2 \qquad \#\{i : X_i > 10\}$$

$$\max(X_1, X_2, \ldots, X_n) \qquad \frac{1}{n}\sum_{i=1}^{n}(X_i - c)^2$$

$$\boxed{\frac{1}{n}\sum_{i=1}^{n} X_i}$$

Many functions of RVs that we care about (**counts**, **means**) involve **sums of RVs**, so we expand on properties of sums of RVs.

Equal vs. Identically Distributed vs. IID

Suppose that we have two random variables X and Y.

X and Y are **equal** if:

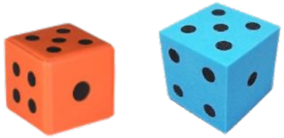
- $X(s) = Y(s)$ for every sample s.
- We write $X = Y$.

X and Y are **identically distributed** if:

- The distribution of X is the same as the distribution of Y
- We say "X and Y are equal in distribution."
- If $X = Y$, then X and Y are identically distributed;
  but the converse is not true (ex: **Y = 7-X**, X is a die)

X and Y are **independent and identically distributed (IID)** if:

- X and Y are identically distributed, *and*
- Knowing the outcome of X does not influence your belief of the outcome of Y, and vice versa ("X and Y are independent.")
- Independence is covered more in STAT 140/EECS 70.
- In CSCI 3022, you will never be expected to prove that RVs are IID.



IID RVs

Have a Normal Day!

[Extra Slides] Derivations

Expected Value of Poisson

https://statproofbook.github.io/P/poiss-mean.html

Properties of Expectation [1/3]

Instead of simulating full distributions, we often just compute expectation and variance directly.

Recall definitions of expectation:

$$\mathbb{E}[X] = \sum_x xP(X = x)$$

Properties:

1. **Expectation is linear**.
Intuition: summations are linear. Proof

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

Properties of Expectation [2/3]

Instead of simulating full distributions, we often just compute expectation and variance directly.

Recall definitions of expectation:

$$\mathbb{E}[X] = \sum_x x P(X = x)$$

$$\mathbb{E}[X] = \sum_{\substack{\text{all samples} \\ s}} X(s)P(s)$$

Properties:

1. **Expectation is linear**.
Intuition: summations are linear. Proof

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

2. Expectation is linear in sums of RVs,
for any relationship between X and Y. Proof

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Properties of Expectation [3/3]

Instead of simulating full distributions, we often just compute expectation and variance directly.

Recall definitions of expectation:

$$\mathbb{E}[X] = \sum_x x P(X = x)$$

$$\mathbb{E}[X] = \sum_{\substack{\text{all samples} \\ s}} X(s)P(s)$$

<u>Properties</u>:

1. **Expectation is linear**.
Intuition: summations are linear. <u>Proof</u>

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

2. Expectation is linear in sums of RVs,
for any relationship between X and Y. <u>Proof</u>

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

3. If g is a non-linear function, then in general $\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$.

- Example: if X is -1 or 1 with equal probability, then $E[X] = 0$ but $E[X^2] = 1 \neq 0$.

There's a more convenient form of variance for use in calculations.

$$\mathbb{V}ar(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

To derive this, we make repeated use of the linearity of expectation.

$$\begin{aligned} Var(X) &= E((X - E(X))^2) \\ &= E\big(X^2 - 2XE(X) + (E(X)^2)\big) \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2 \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

Variance: An Alternate Derivation

Jump back: link

There's a more convenient form of variance for use in calculations.

$$\mathbb{V}ar(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

To derive this, we make repeated use of the linearity of expectation.

$$\begin{aligned} Var(X) &= E((X - E(X))^2) \\ &= E\big(X^2 - 2XE(X) + (E(X)^2)\big) \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2 \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

Properties of Expectation #1

Jump back: link

Recall definition of expectation:

$$\mathbb{E}[X] = \sum_x x P(X = x)$$

$$\mathbb{E}[X] = \sum_{\substack{\text{all samples} \\ s}} X(s)P(s)$$

1. **Expectation is linear**:

(intuition: summations are linear)

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

Proof:

$$\begin{aligned}
\mathbb{E}[aX + b] &= \sum_x (ax + b)P(X = x) = \sum_x \big(axP(X = x) + bP(X = x)\big) \\
&= a\sum_x xP(X = x) + b\sum_x P(X = x) \\
&= a\mathbb{E}[X] + b \cdot 1
\end{aligned}$$

Properties of Expectation #2

Jump back: link

Recall definitions of expectation: $\mathbb{E}[X] = \sum_x x P(X = x)$ $\mathbb{E}[X] = \sum_{\substack{\text{all samples} \\ s}} X(s)P(s)$

3. **Expectation is linear in sums of RVs**:

For any relationship between X and Y.

Proof:

$$\begin{aligned}
\mathbb{E}[X+Y] &= \sum_s (X+Y)(s)P(s) = \sum_s \big(X(s)+Y(s)\big)P(s) \\
&= \sum_s \big(X(s)P(s) + Y(s)P(s)\big) \\
&= \sum_s X(s)\,P(s) + \sum_s Y(s)P(s) \\
&= \mathbb{E}[X] + \mathbb{E}[Y]
\end{aligned}$$

Properties of Variance #1

Jump back: link

We know that $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$

In order to compute Var(aX + b), consider:

- A shift by **b** units **does not** affect spread. Thus, Var(aX + b) = Var(aX).
- The multiplication by **a does** affect spread!

Then,

$$\begin{aligned} Var(aX+b) = Var(aX) &= E((aX)^2) - (E(aX))^2 \\ &= E(a^2X^2) - (aE(X))^2 \\ &= a^2\left(E(X^2) - (E(X))^2\right) \\ &= a^2Var(X) \end{aligned}$$

In summary:

$$\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$$

$$\mathbb{V}ar(aX+b) = a^2\mathbb{V}ar(X)$$
$$\mathbb{SD}(aX+b) = |a|\mathbb{SD}(X)$$

Don't forget the absolute values and squares!

Properties of Variance #2

Jump back: link

The variance of a sum is affected by the dependence between the two random variables that are being added. Let's expand out the definition of Var(X + Y) to see what's going on.

Let $\mu_x = E[X], \mu_y = E[Y]$

By the linearity of expectation, and the substitution.

We see

$$
\begin{aligned}
Var(X+Y) &= E\big[(X+Y-E(X+Y))^2\big] \\
&= E\big[((X-\mu_x)+(Y-\mu_y))^2\big] \\
&= E\big[(X-\mu_x)^2 + 2(X-\mu_x)(Y-\mu_y) + (Y-\mu_y)^2\big] \\
&= E\big[(X-\mu_x)^2\big] + E\big[(Y-\mu_y)^2\big] + 2E\big[(X-\mu_x)(Y-\mu_y)\big] \\
&= Var(X) + Var(Y) + 2E\big[(X-E(X))(Y-E(Y))\big]
\end{aligned}
$$

Addition rule for variance

If X and Y are **uncorrelated** (in particular, if they are **independent**), then

$$\mathbb{V}ar(X+Y) \;=\; \mathbb{V}ar(X) + \mathbb{V}ar(Y)$$

Therefore, under the same conditions,

$$\mathbb{SD}(X+Y) \;=\; \sqrt{\mathbb{V}ar(X) + \mathbb{V}ar(Y)} \;=\; \sqrt{(\mathbb{SD}(X))^2 + (\mathbb{SD}(Y))^2}$$

- Think of this as “Pythagorean theorem” for random variables.
- Uncorrelated random variables are like orthogonal vectors.

Properties of Variance #1

We know that $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$

In order to compute Var(aX + b), consider:

- A shift by **b** units **does not** affect spread. Thus, Var(aX + b) = Var(aX).
- The multiplication by **a does** affect spread!

Then,

$$\begin{aligned} Var(aX+b) = Var(aX) &= E((aX)^2) - (E(aX))^2 \\ &= E(a^2X^2) - (aE(X))^2 \\ &= a^2\left(E(X^2) - (E(X))^2\right) \\ &= a^2Var(X) \end{aligned}$$

In summary:

$$\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$$

$$\mathbb{V}ar(aX+b) = a^2\mathbb{V}ar(X)$$
$$\mathbb{SD}(aX+b) = |a|\mathbb{SD}(X)$$

Don't forget the absolute values and squares!

Properties of Variance #2

Jump back: link

The variance of a sum is affected by the dependence between the two random variables that are being added. Let's expand out the definition of Var(X + Y) to see what's going on.

Let $\mu_x = E[X], \mu_y = E[Y]$

By the linearity of expectation, and the substitution.

We see

$$\begin{aligned} Var(X+Y) &= E\left[(X+Y-E(X+Y))^2\right] \\ &= E\left[((X-\mu_x)+(Y-\mu_y))^2\right] \\ &= E\left[(X-\mu_x)^2+2(X-\mu_x)(Y-\mu_y)+(Y-\mu_y)^2\right] \\ &= E\left[(X-\mu_x)^2\right]+E\left[(Y-\mu_y)^2\right]+2E\left[(X-\mu_x)(Y-\mu_y)\right] \\ &= Var(X)+Var(Y)+2E\left[(X-E(X))(Y-E(Y))\right] \end{aligned}$$

Addition rule for variance

If X and Y are **uncorrelated** (in particular, if they are **independent**), then

$$\mathbb{V}ar(X+Y) \;=\; \mathbb{V}ar(X) + \mathbb{V}ar(Y)$$

Therefore, under the same conditions,

$$\mathbb{SD}(X+Y) \;=\; \sqrt{\mathbb{V}ar(X) + \mathbb{V}ar(Y)} \;=\; \sqrt{(\mathbb{SD}(X))^2 + (\mathbb{SD}(Y))^2}$$

- Think of this as “Pythagorean theorem” for random variables.
- Uncorrelated random variables are like orthogonal vectors.

LECTURE 14

Expected Value & Variance

Narges Norouzi and Lisa Yan

Content credit: Acknowledgments