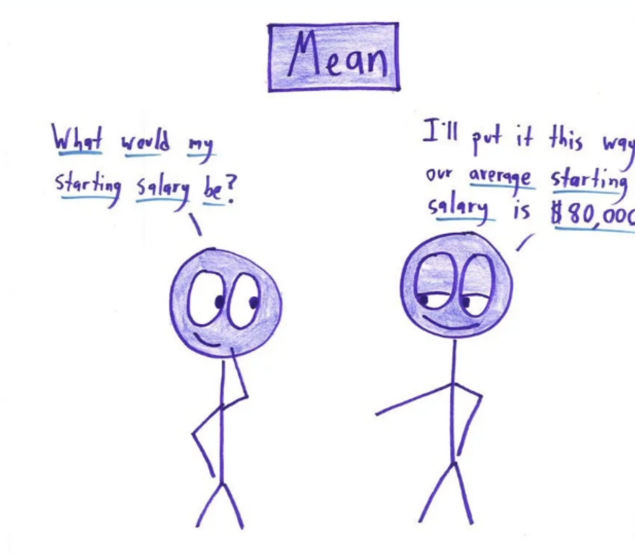
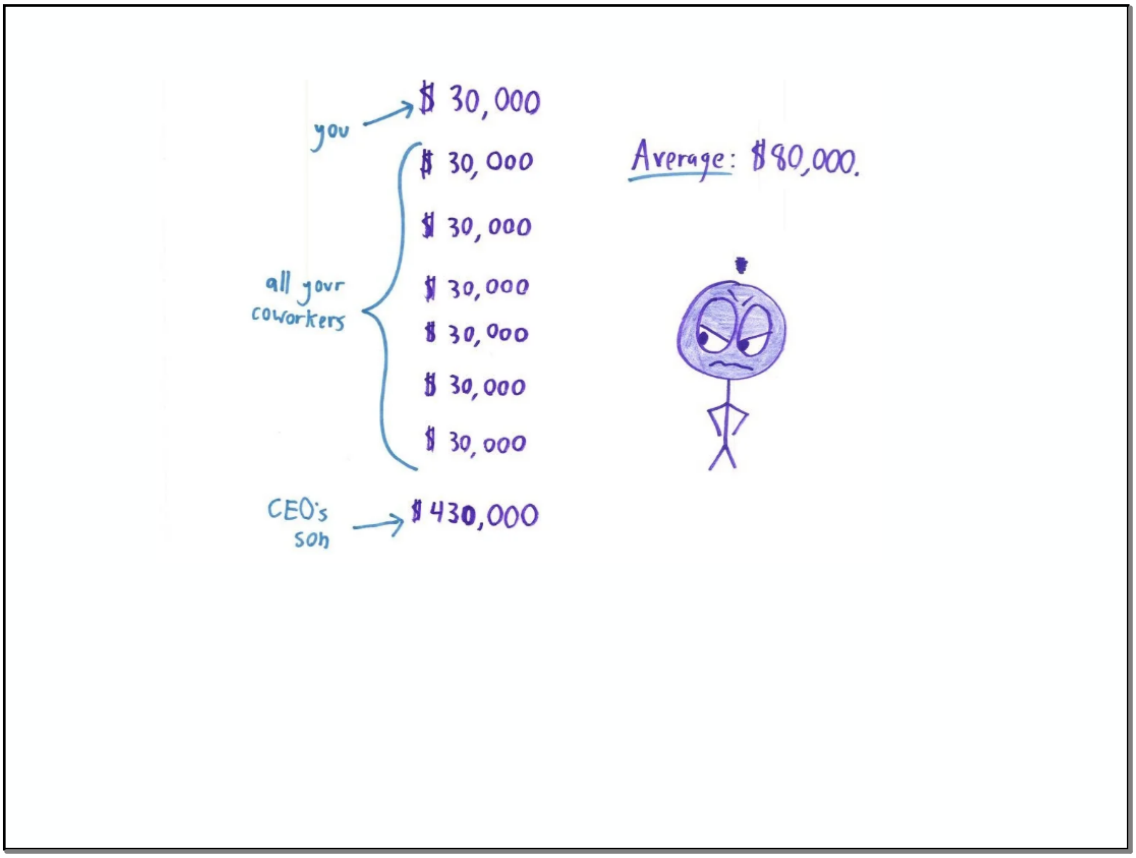
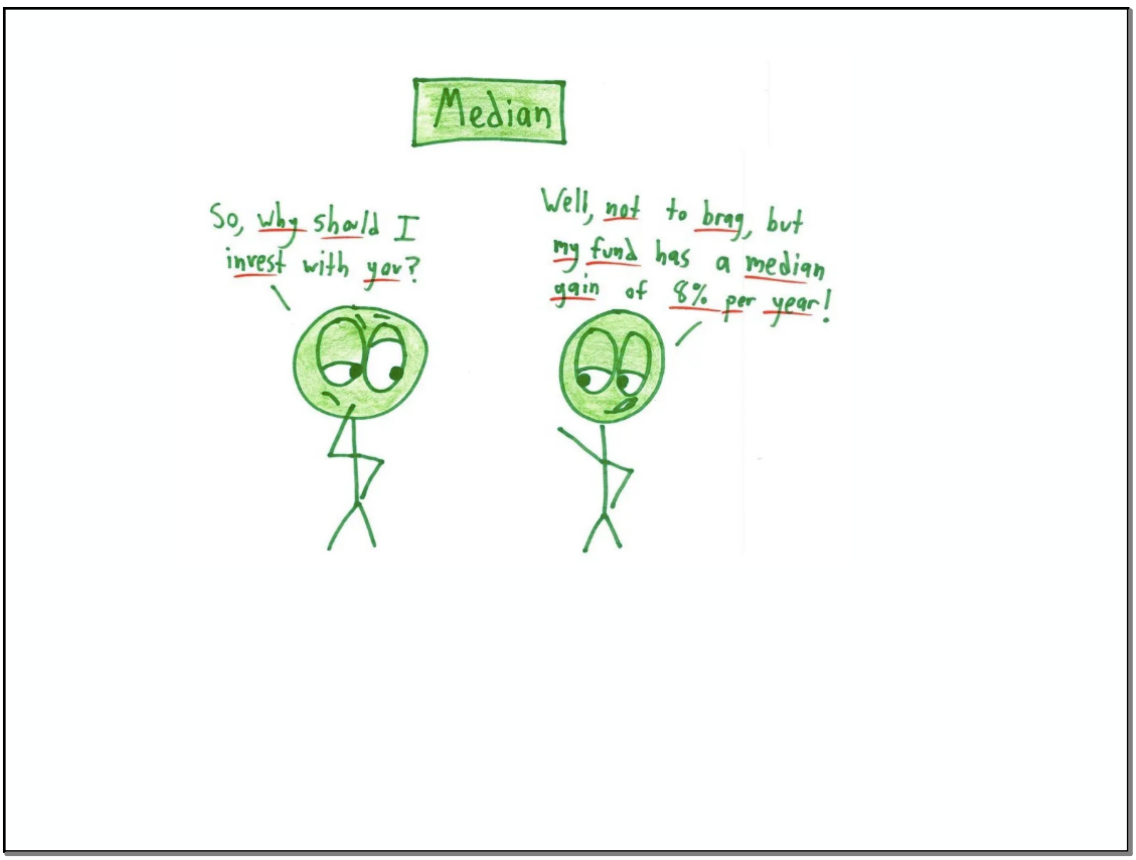
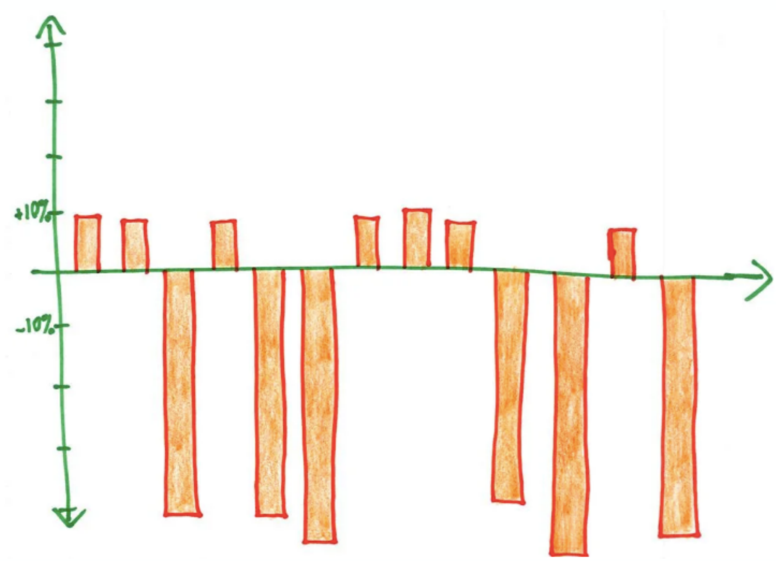
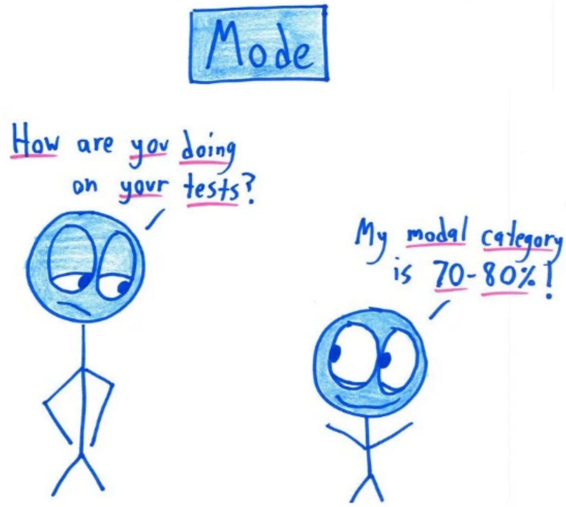

Mean
What would my starting salary be?
I'll put it this way: our average starting salary is $80,000!

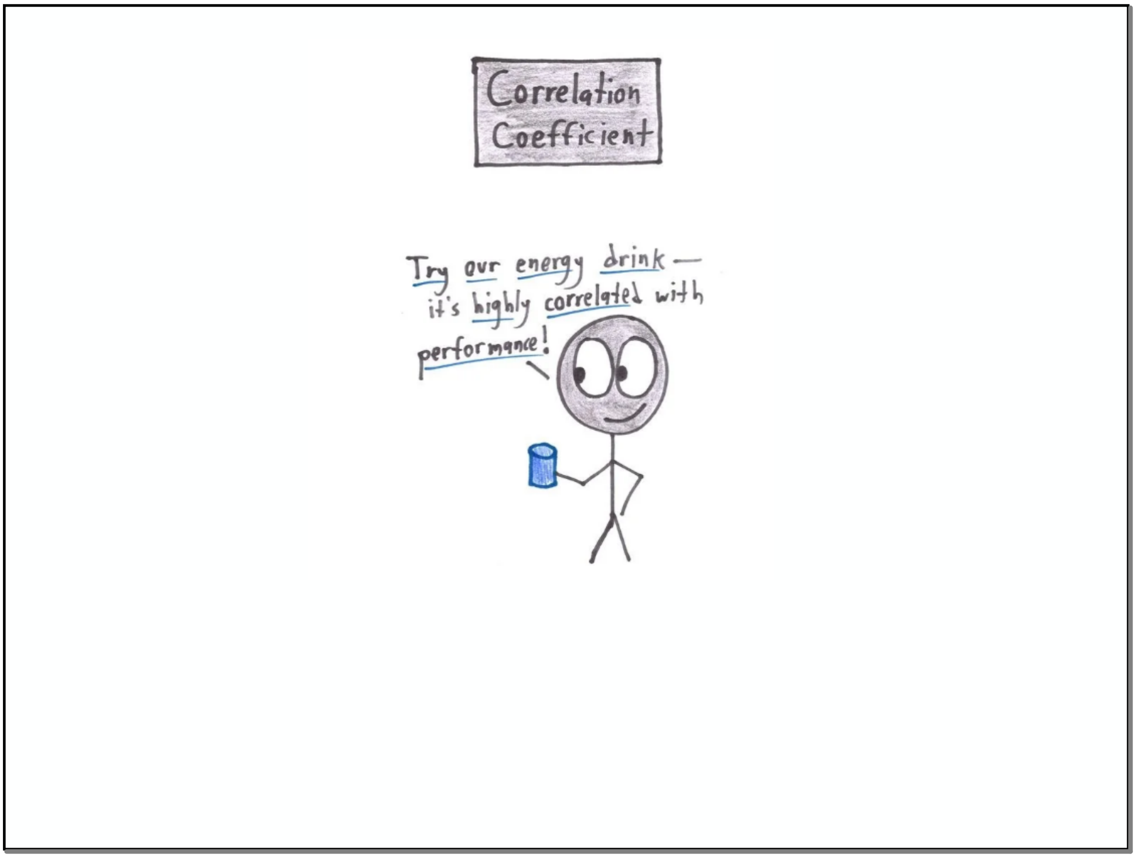

you
$ 30,000
$ 30,000
$ 30,000
$ 30,000
$ 30,000
$ 30,000
$ 30,000
all your coworkers
Average: $80,000.
CEO's son
$430,000

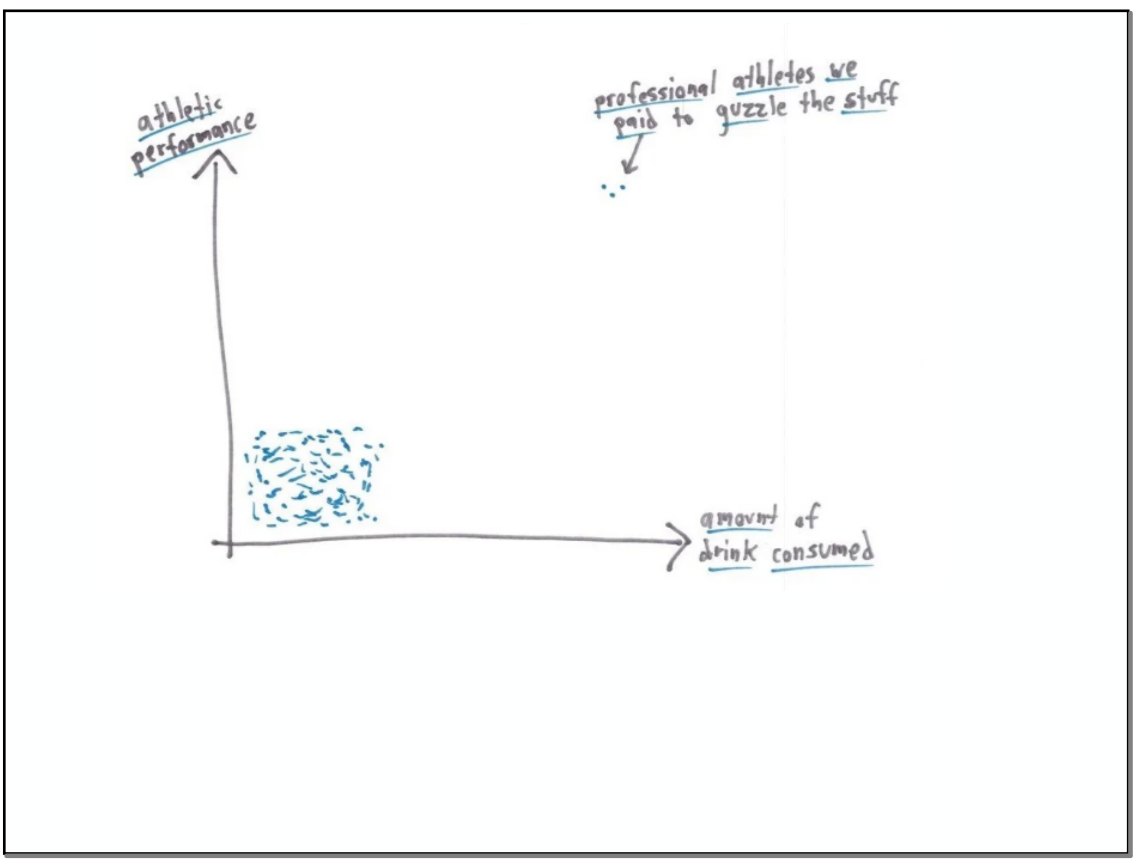

Median
So, why should I invest with you?
Well, not to brag, but my fund has a median gain of 8% per year!

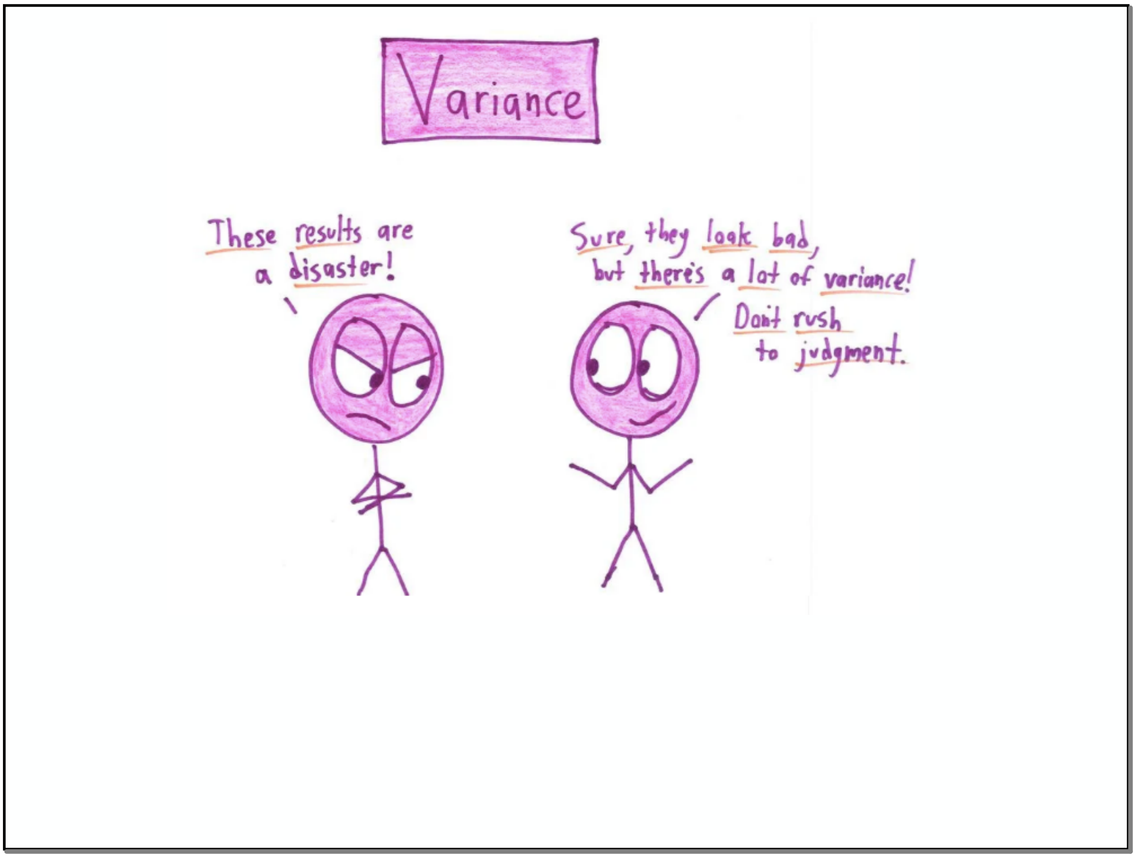

+10%
-10%

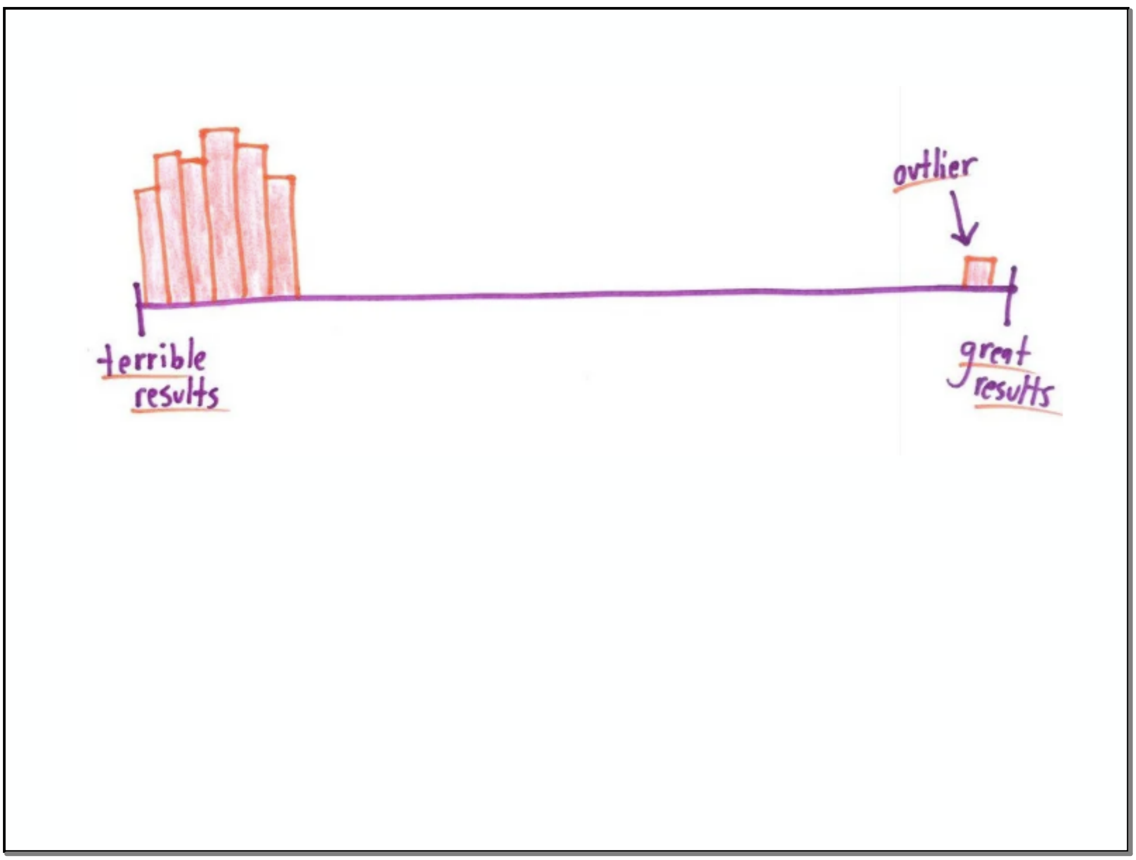


Score Category	Number of Tests
90s	0
80s	0
70s	2
60s	1
50s	1
40s	1
30s	1
20s	1


Correlation Coefficient
Try our energy drink — it's highly correlated with performance!

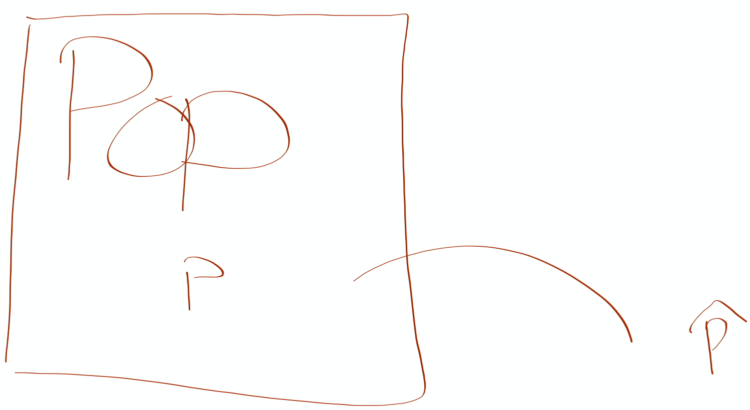

athletic performance
professional athletes we paid to guzzle the stuff
amount of drink consumed


Variance
These results are a disaster!
Sure, they look bad, but there's a lot of variance! Don't rush to judgment.


outlier
terrible results
great results

By the end of this section, you should be able to:

- ✓ CALCULATE the mean and standard deviation of the sampling distribution of a sample proportion $\hat{p}$ and INTERPRET the standard deviation.
- ✓ DETERMINE if the sampling distribution of $\hat{p}$ is approximately Normal.
- ✓ If appropriate, USE a Normal distribution to CALCULATE probabilities involving $\hat{p}$.



Lesson 7.2: What's the proportion of orange Reese's Pieces?

If we take a sample of Reese's Pieces, what proportion of the candies will be orange?

Suppose a large bag of Reese's Pieces has 1000 pieces. The manufacturer says that exactly 40% of the candies are orange. If we select a sample of 50 pieces, how many will be orange?
Let X = the number of orange candies in the sample.

1. What type of probability distribution does X have? Justify.

B I N S

Lesson 7.2: What's the proportion of orange Reese's Pieces?

If we take a sample of Reese's Pieces, what proportion of the candies will be orange?

Suppose a large bag of Reese's Pieces has 1000 pieces. The manufacturer says that exactly 40% of the candies are orange. If we select a sample of 50 pieces, how many will be orange?
Let X = the number of orange candies in the sample.

1. What type of probability distribution does X have? Justify.

If we take a sample of Reese's Pieces, what proportion of the candies will be orange?

Suppose a large bag of Reese's Pieces has 1000 pieces. The manufacturer says that exactly 40% of the candies are orange. If we select a sample of 50 pieces, how many will be orange?
Let X = the number of orange candies in the sample.

1. What type of probability distribution does X have? Justify.

$\boxed{B}$ Binary
- Sucess → orange
- failure → not orange

$\boxed{I}$ 10% condition
50 < 100
$\le \frac{1}{10} \cdot 1000$

$\boxed{N}$ n = 50

$\boxed{S}$ p = .40

2. Draw a sample of 50 Reese's Pieces using the **Rossman/Chance Applet Collection (Google it).** applet. How many pieces were orange? Repeat this 5 times. Write the values below.

$X =$ $X =$ $X =$ $X =$ $X =$

Google
Rossman/Chance Applet

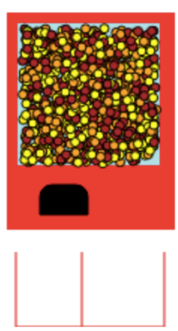
- Least Squares Regression (js)

Sampling Distribution Simulations

- Reeses Pieces (js)
- Sampling Words (js)
- Sampling from a Finite Population/Model/Bootstrap (js)
- Simulating Confidence Intervals for Population Parameter (js)
- Improved Batting Averages (Power) (js)
- ANOVA simulation (js)
- NEW: Guess the p-value (js)

Reeses Pieces

Probability of orange 0.5
Number of candies 25
Number of samples 1
☑ Animate
Draw Samples
Total = 0



◉ Number of orange
○ Proportion of orange

As extreme as ≥

Count

☐ Summary Stats

3. Write the values on sticker dots and add it to the dotplot on the board. Sketch the dotplot below.

4. What does each dot represent?

\# orange pieces in a sample of 50

5. What is the **mean** and the **standard deviation** for this binomial distribution of X? Show work.

$$\mu_x = np = (50)(.4) =$$

(binomial)

$$\sigma_x = \sqrt{np(1-p)}$$

(binomial)

4. What does each dot represent?

The number of orange pieces from a sample of 50

5. What is the **mean** and the **standard deviation** for this binomial distribution of X? Show work.

$$\mu_x = np = 50 \times 0.4 = \underline{20} \qquad \sigma_x = \sqrt{np(1-p)} = \sqrt{50 \cdot 0.4(.6)} = \underline{3.46}$$

6. What is the approximate shape of the sampling distribution for X? Explain and sketch it below.

Large Counts Condition

- $np \geq 10$ $50(.4) = 20$ ✓
  $n(1-p) \geq 10$ $50(.6) = 30$ ✓

So distribution is approxim. Normal.

$N(20, 3.46)$

Instead of finding the number of candies that are orange, we will now find the proportion of candies that are orange.

7. Use your samples from #2 and turn each number of orange candies into the **proportion of orange candies** in the sample ($\hat{p}$). Write the proportions below and add them to the second dotplot on the board.

$\hat{p} = \frac{29}{50} = .58$ $\hat{p} =$ $\hat{p} =$ $\hat{p} =$ $\hat{p} =$

$X = 29$

.2 .22 ●.24 .26 .28

.3 ●.32 .34 ●.36 .38

●.4 ●.42 .44 .46 .48

8. Sketch the dotplot using your $\hat{p}$ (sample proportion values) below.

9. What does each dot represent?

The proportion of orange pieces from a sample of 50

10. Find the new mean and standard deviation. Show work.

divide by 50

$\mu_{\hat{p}} = \frac{20}{50} = 0.40$

SD $\sigma_{\hat{p}} = \frac{\sqrt{np(1-p)}}{50} = \frac{\sqrt{50(.4)(.6)}}{50} = .069$

$\mu_{\hat{p}} = \frac{np}{n} = p$

$\sigma_{\hat{p}} =$

$$\sqrt{25x}$$

$$= 5\sqrt{x}$$

$$7\sqrt{x}$$

$$= \sqrt{49x}$$

$$\sqrt{\frac{x}{n}}$$

$$= \sqrt{\frac{x}{4n}}$$

$$= \frac{1}{2}\sqrt{\frac{x}{n}}$$

VARIABILITY

$\underset{n}{1000} \rightarrow \underset{n}{9000}$

$$\sqrt{\frac{p(1-p)}{n}}$$

$$= \sqrt{\frac{p(1-p}{9n}}$$

$$= \frac{1}{3}\sqrt{\frac{p(1-p)}{n}}$$

10. Find the new mean and standard deviation. Show work.

divide by 50

$\mu_{\hat{p}} = \frac{20}{50} = 0.40$

SD $\sigma_{\hat{p}} = \frac{\sqrt{np(1-p)}}{50} = \frac{\sqrt{50(.4)(.6)}}{50} = .069$

$\mu_{\hat{p}} = \frac{np}{n} = p$

$\sigma_{\hat{p}} = \frac{\sqrt{np(1-p)}}{n} = \sqrt{\frac{\cancel{n}p(1-p)}{n^{\cancel{2}}}} = \sqrt{\frac{p(1-p)}{n}}$

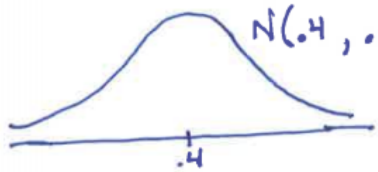
$\downarrow$

$\frac{1}{n}\sqrt{np(1-p)}$

11. What is the approximate shape of the sampling distribution for $\hat{p}$? Explain and sketch it below.

Normal because of Large Counts

$N(.4, .069)$

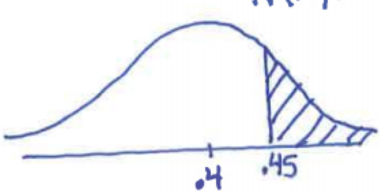


$50(.4) = 20 \checkmark$

$50(.6) = 30 \checkmark$

12. We know that bags of Reese's Pieces contain exactly 40% that are orange. If we select a random sample of 50 candies, what is the probability that the sample proportion will be ~~50~~ 45% or greater?

$N(.4, .069)$



$$Z = \frac{\hat{p} - p}{\sigma}$$

$$= \frac{.45 - .40}{.069}$$

$$= .72$$

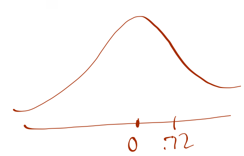


TABLE A

$\boxed{= .2358}$

$P(z > .72) = \text{normalcdf}[.72, 1000, 0, 1]$ (L U μ σ)

The Sampling Distribution of $\hat{p}$ (sample proportions)

Important ideas:

The Sampling Distribution of $\hat{p}$ (sample proportions)

Important ideas:

Mean

$\mu_{\hat{p}} = p$

Std. Dev.

$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

if 10% CONDITION MET

$n < .10N$

Normal Approx.

if Large Counts

$np \geq 10$

$n(1-p) \geq 10$

for Probability if sampling distrib. of $\hat{p}$ is approx. normal.

$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$

on-line
videos

Suppose that 75% of young adult Internet users (ages 18 to 29) watch online videos. A polling organization contacts an SRS of 1000 young adult Internet users and calculates the proportion $\hat{p}$ in this sample who watch online videos.

1. Identify the mean of the *sampling distribution* of $\hat{p}$.

2. Calculate and interpret the standard deviation of the sampling distribution of $\hat{p}$. Check that the 10% condition is met.

1. Identify the mean of the *sampling distribution* of $\hat{p}$. $\mu_{\hat{p}} = p = 0.75$

2. Calculate and interpret the standard deviation of the sampling distribution of $\hat{p}$. Check that the 10% condition is met.

1000 < 10% of all internet users, (18-29)

$\sigma_{\hat{p}} = \sqrt{\frac{.75(.25)}{1000}} = .014$

The proportion of young adults who watch online videos in a sample of 1000 typically varies by 0.014 from the true proportion of 0.75

3. Is the sampling distribution of $\hat{p}$ approximately Normal? Check that the Large Counts condition is met.

4. Find the probability that the random sample of 1000 young adults will give a result within 2 percentage points of the true value.

3. Is the sampling distribution of $\hat{p}$ approximately Normal? Check that the Large Counts condition is met.

$(.75)(1000) = 750$
$(.25)(1000) = 250$ ≥ 10 so, yes, it is approx. normal.

4. Find the probability that the random sample of 1000 young adults will give a result within 2 percentage points of the true value.

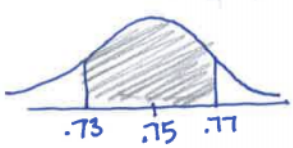
3. Is the sampling distribution of $\hat{p}$ approximately Normal? Check that the Large Counts condition is met.

$(.75)(1000) = 750$
$(.25)(1000) = 250$ ≥ 10 so, yes, it is approx. normal.

4. Find the probability that the random sample of 1000 young adults will give a result within 2 percentage points of the true value.

$N(.75, .014)$

$Z = \frac{.73-.75}{.014} = -1.43$ $Z = \frac{.77-.75}{.014} = 1.43$



$P(-1.43 < z < 1.43)$

$= \text{normalcdf}[\underset{\text{Low}}{-1.43}, \underset{\text{Upper}}{1.43}, \underset{\mu}{0}, \underset{\sigma}{1}] = 0.8472$

5. If the sample size were 9000 rather than 1000, how would this change the sampling distribution of $\hat{p}$?

5. If the sample size were 9000 rather than 1000, how would this change the sampling distribution of $\hat{p}$?

- The shape would remain approx. normal.
- The center would stay the same (.75)
- the variability woud decrease by $\frac{1}{3}$

LCQ

7.2....35, 37, 41, 43, 47–50

and study pp. 458–465