Sampling Distribution of a Sample Proportion

Lecture 26
Section 8.4

Robb T. Koether

Hampden-Sydney College

Mon, Mar 5, 2012

Outline

Outline

Sampling Distributions

Definition (Sampling Distribution of a Statistic)

The sampling distribution of a statistic is the distribution of values of that statistic over all possible samples of a given size n from the population.

- We may sample with or without replacement.
- For our purposes, it will be simpler to sample with replacement.

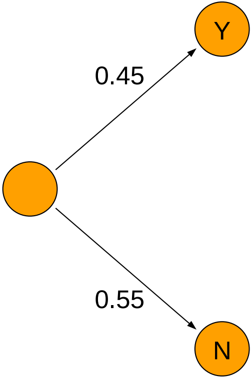
The Sample Proportion

- We will work out the sampling distribution for $\hat{p}$ for sample sizes of 1, 2, and 3.
- Then I will show you the sampling distribution for $\hat{p}$ for sample sizes of 4, 5, and 10.

Example

- Suppose that 45% of all people approve of President Obama's performance.
- Suppose that we select one person at random.
- We may diagram the 2 possibilities.

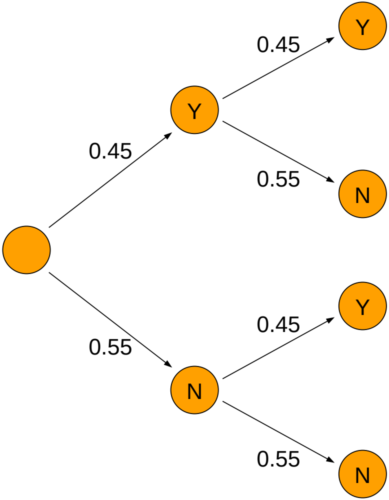
Example



Example

- Now we take a sample of 2 people, sampling with replacement.
- Find the sampling distribution of $\hat{p}$.

Example



YY	2	$(0.45)^2 = 0.2025$
YN	1	$(0.45)(0.55) = 0.2475$
NY	1	$(0.55)(0.45) = 0.2475$
NN	0	$(0.55)^2 = 0.3025$

Example

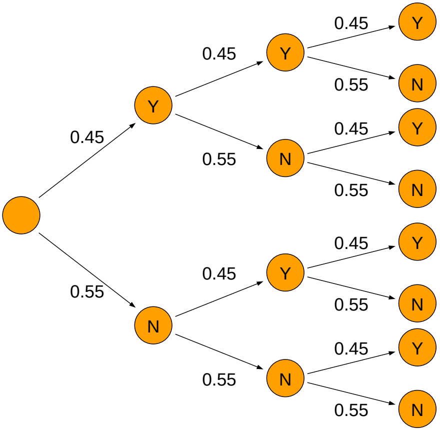
- Let x be the number of people (out of 2) who strong disapprove of President Obama's performance.
- The probability distribution of $\hat{p}$ is

$\hat{p}$	$P(\hat{p})$
0	0.2025
1/2	0.4950
1	0.3025

Example

- Now we take a sample of 3 people, sampling with replacement.
- Find the sampling distribution of $\hat{p}$.

Example



YYY	3	$(0.45)^3 = 0.0911$
YYN	2	$(0.45)^2(0.55) = 0.1114$
YNY	2	$(0.45)(0.55)(0.45) = 0.1114$
YNN	1	$(0.45)(0.55)^2 = 0.1361$
NYY	2	$(0.55)(0.45)^2 = 0.1114$
NYN	1	$(0.55)(0.45)(0.55) = 0.1361$
NNY	1	$(0.55)^2(0.45) = 0.1114$
NNN	0	$(0.55)^3 = 0.1664$

Example

- Let $\hat{p}$ be the sample proportion of people who strong disapprove of President Obama's performance.
- The sampling distribution of $\hat{p}$ is

$\hat{p}$	$P(\hat{p})$
0	0.1664
1/3	0.4084
2/3	0.3341
1	0.0911

Samples of Size $n = 4$

- If we sample 4 people, then the sampling distribution of the sample proportion is

$\hat{p}$	$P(\hat{p})$
0	0.0915
$1/4$	0.2995
$2/4$	0.3675
$3/4$	0.2005
1	0.0410

Samples of Size $n = 5$

- If we sample 5 people, then the sampling distribution of the sample proportion is

$\hat{p}$	$P(\hat{p})$
0	0.0503
1/5	0.2059
2/5	0.3369
3/5	0.2757
4/5	0.1128
1	0.0185

Samples of Size $n = 6$

- If we sample 6 people, then the sampling distribution of the sample proportion is

$\hat{p}$	$P(\hat{p})$
0	0.0277
$1/6$	0.1359
$2/6$	0.2780
$3/6$	0.3032
$4/6$	0.1861
$5/6$	0.0609
1	0.0083

Samples of Size $n = 10$

- If we sample 10 people, then the sampling distribution of the sample proportion is

$\hat{p}$	$P(\hat{p})$
0.00	0.0025
0.10	0.0207
0.20	0.0763
0.30	0.1665
0.40	0.2384
0.50	0.2340
0.60	0.1596
0.70	0.0746
0.80	0.0229
0.90	0.0042
1.00	0.0003

Outline

The Central Limit Theorem for Proportions

Theorem (The Central Limit Theorem for Proportions)

- *For any population and any sample size, the sampling distribution of $\hat{p}$ has the following mean and standard deviation:*

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

- *Furthermore, the sampling distribution of $\hat{p}$ is approximately normal, provided n is large enough.*

The Central Limit Theorem for Proportions

The Sample Size

- n is large enough if

$$np \geq 5 \text{ and } n(1 - p) \geq 5.$$

- If n is small, then we have to work out the distribution by hand.

Outline

Applications

- Suppose that 60% of all high-school students own a cell phone.
- If we survey 3 high-school students, how likely is it that we will find that at least 2 of them own a cell phone?

Applications

- Suppose that 60% of all high-school students own a cell phone.
- If we survey 150 high-school students, how likely is it that we will find that at least 65% of them own a cell phone?

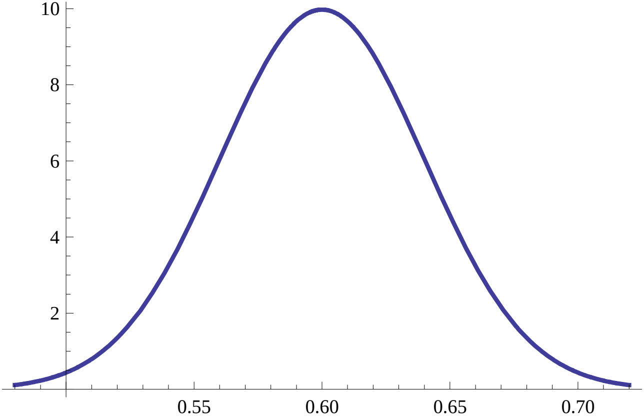
Applications

- If $p = 0.60$ and our sample size is $n = 150$, then $\hat{p}$ is normal with mean $\mu_{\hat{p}} = 0.60$ and

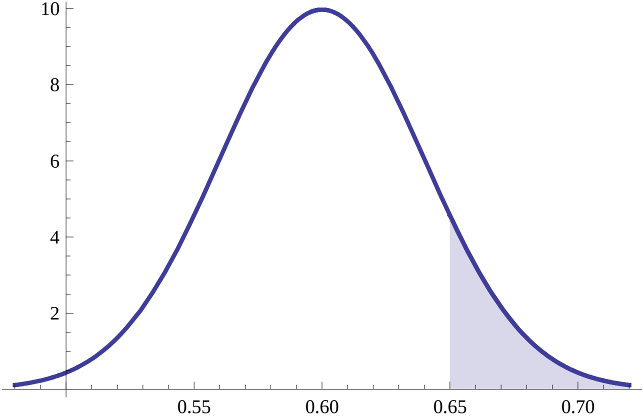
$$\sigma_{\hat{p}} = \sqrt{\frac{(0.60)(0.40)}{150}} = \sqrt{0.0016} = 0.04.$$

- We want to know the probability that $\hat{p} \geq 0.65$.

Applications



Applications



Applications

- The probability that $\hat{p}$ is greater than 0.65 is

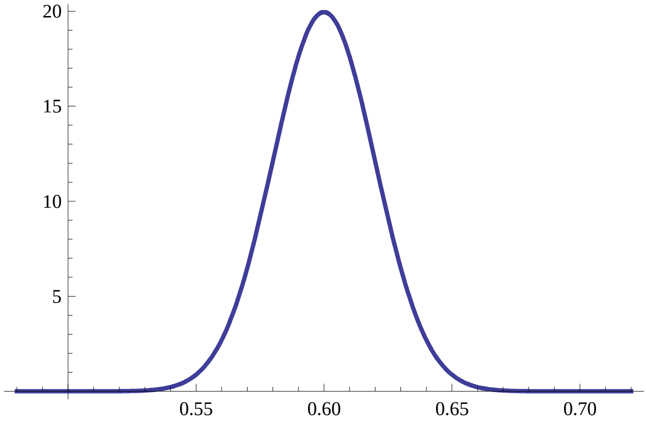
$$\texttt{normalcdf(.65,E99,.60,.04)} = 0.1056.$$

Applications

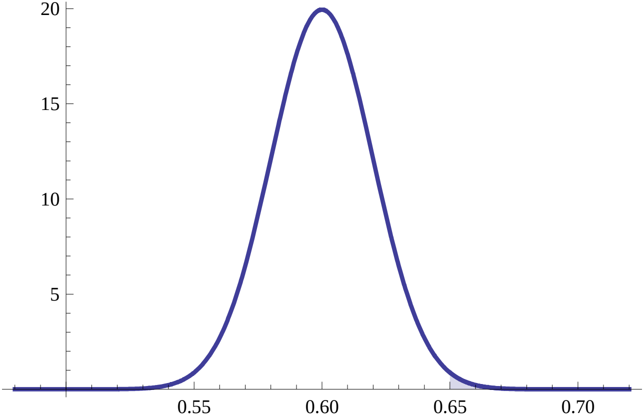
- What if our sample size were 600 instead of 150?
- Then $\hat{p}$ is normal with mean $\mu_{\hat{p}} = 0.60$ and

$$\sigma_{\hat{p}} = \sqrt{\frac{(0.60)(0.40)}{600}} = \sqrt{0.0004} = 0.02.$$

Applications



Applications



- The probability that $\hat{p}$ is greater than 0.65 is

$$\texttt{normalcdf(.65,E99,.60,.02)} = 0.0062.$$

Guessing on a Test

Example (Guessing on a Test)

- A student takes a math placement test with 25 multiple-choice questions.

Guessing on a Test

Example (Guessing on a Test)

- A student takes a math placement test with 25 multiple-choice questions.
- Each question has 5 choices.

Guessing on a Test

Example (Guessing on a Test)

- A student takes a math placement test with 25 multiple-choice questions.
- Each question has 5 choices.
- If a student guesses at each answer, what will his score most likely be?

Guessing on a Test

Example (Guessing on a Test)

- A student takes a math placement test with 25 multiple-choice questions.
- Each question has 5 choices.
- If a student guesses at each answer, what will his score most likely be?
- If a student scores 10 out of 25, can we be sure that he did not guess at all 25 answers?

Guessing on a Test

Example (Guessing on a Test)

- If he guesses at all 25 answers, then the probability of getting any one answer correct is $p = 0.20$.
- The proportion that he actually gets correct is $\hat{p}$.

Guessing on a Test

Example (Guessing on a Test)

- If he guesses at all 25 answers, then the probability of getting any one answer correct is $p = 0.20$.
- The proportion that he actually gets correct is $\hat{p}$.
- The distribution of $\hat{p}$ is normal with mean 0.20 and standard deviation $\sqrt{\frac{(0.20)(0.80)}{25}} = 0.04$.

Guessing on a Test

Example (Guessing on a Test)

- If he guesses at all 25 answers, then the probability of getting any one answer correct is $p = 0.20$.
- The proportion that he actually gets correct is $\hat{p}$.
- The distribution of $\hat{p}$ is normal with mean 0.20 and standard deviation $\sqrt{\frac{(0.20)(0.80)}{25}} = 0.04$.
- If we allow three standard deviations each way, then $\hat{p}$ is almost certainly at least 0.08 and at most 0.32.

Guessing on a Test

Example (Guessing on a Test)

- If he guesses at all 25 answers, then the probability of getting any one answer correct is $p = 0.20$.
- The proportion that he actually gets correct is $\hat{p}$.
- The distribution of $\hat{p}$ is normal with mean 0.20 and standard deviation $\sqrt{\frac{(0.20)(0.80)}{25}} = 0.04$.
- If we allow three standard deviations each way, then $\hat{p}$ is almost certainly at least 0.08 and at most 0.32.
- That represents from 2 to 8 correct answers.

Outline

Assignment

Homework

- Read Sections 8.1 - 8.2, pages 499 - 508.
- Exercises 7 - 14, page 526.