AP Stats
More Examples 7.2

Name Key
Date Hour

Sample Proportions

Summary

In this section, we learned that...

When we want information about the population proportion p of successes, we often take an SRS and use the sample proportion $\hat{p}$ to estimate the unknown parameter p. The **sampling distribution** of $\hat{p}$ describes how the statistic varies in all possible samples from the population.

① Random Condition

The **mean** of the sampling distribution of $\hat{p}$ is equal to the population proportion p. That is, $\hat{p}$ is an unbiased estimator of p.

The **standard deviation** of the sampling distribution of $\hat{p}$ is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ for an SRS of size n. This formula can be used if the population is at least 10 times as large as the sample (the 10% condition). The standard deviation of $\hat{p}$ gets smaller as the sample size n gets larger.

② 10 % condition

When the sample size n is larger, the sampling distribution of $\hat{p}$ is close to a Normal distribution with mean p and standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.

In practice, use this Normal approximation when both $np \geq 10$ and $n(1 - p) \geq 10$ (the Normal condition).

③ Normal Condition

1) A polling organization asks an SRS of 1500 first-year college students how far away their home is. Suppose that 35% of all first-year students actually attend college within 50 miles of home. What is the probability that the random sample of 1500 students will give a result within 2 percentage points of this true value?

State: What is probability that $\hat{p}$ falls between .33 and .37?

$P(.33 \leq \hat{p} \leq .37)$

Plan: ① Have SRS of size $n = 1500$, in which $p = .35$

so $\mu_{\hat{p}} = p = .35$

② ✓ 10 % condition for $\sigma_{\hat{p}}$

$10(1500) \leq$ Population

$15000 \leq$ Population

Reasonable to assume pop. 1st year college students is at least 15,000

$\sigma_{\hat{p}} = \sqrt{\frac{(.35)(.65)}{1500}} \approx .0123$

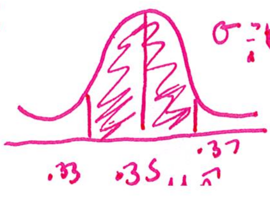
③ Is Dist approximately Normal?

$1500(.35) \geq 10$ and $1500(.65) \geq 10$

$525 \geq 10$ ✓ $975 \geq 10$ ✓

so distribution is $\approx$ Normal with $N(.35, .0123)$

DO:



$\sigma \approx .0123$

$P(.33 \leq \hat{p} \leq .37) = \text{normcdf}(.33, .37, .35, .0123) \approx .8961$

Conclude: About 90% of all SRS of size 1500 will give a result within 2 percentage points of the truth about the population.

2) Harley-Davidson motorcycles make up 14% of all the motorcycles registered in the United States. You plan to interview an SRS of 500 motorcycle owners. How likely is your sample to contain 20% or more who own Harleys? Follow 4 Step Process.

State: $P(\hat{p} \geq .20)$?

Plan: ① Have SRS size $n = 500$, where $p = .14$, so $\mu_{\hat{p}} = p = .14$.

② ✓ 10% for $\sigma_{\hat{p}}$

$10(500) \leq$ population of motorcycles

$5000 \leq$ pop of motorcycles

Reasonable to assume

so $\sigma_{\hat{p}} = \sqrt{\frac{(.14)(.86)}{500}} = .0155$

③ ✓ if Dist is ≈ Normal

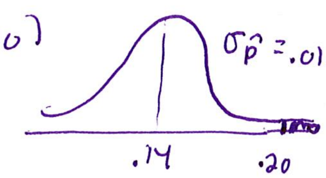
$500(.14) \geq 10$ and $500(.86) \geq 10$

$70 \geq 10$ ✓ $430 \geq 10$ ✓

so shape is approximately Normal

$N(.14, .0155)$

Do: $P(\hat{p} \geq .20)$



normcdf$(.20, \infty, .14, .0155) \approx 5.4 \times 10^{-5} \approx 0$

Conclude: While it is possible, it is extremely unlikely that we would get a sample of 500 motorcycles in which at least 20% are Harleys.

3) A sample survey interviews an SRS of 267 college women. Suppose (as is roughly true) that 70% of college women have been on a diet within the past 12 months. What is the probability that 75% or more of the women in the sample have been on a diet. Follow 4 Step process.

State: $P(\hat{p} \geq .75)$?

Plan: ① Have SRS size $n = 267$, where $p = .70$, so $\mu_{\hat{p}} = p = .70$

② ✓ 10% for $\sigma_{\hat{p}}$

$10(267) \geq$ pop of college women

$2{,}670 \geq$ pop.

reasonable to assume

so $\sigma_{\hat{p}} = \sqrt{\frac{(.70)(.30)}{267}} \approx .0280$

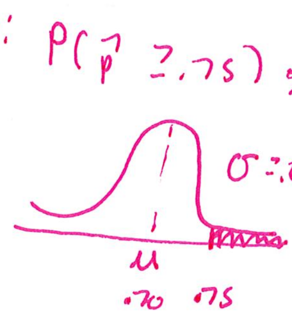
③ ✓ if Dist ≈ Normal

$267(.70) \geq 10$ and $267(.30) \geq 10$

$186.9 \geq 10$ ✓ $80.1 \geq 10$ ✓

$\therefore \sim N(.70, .0280)$

Do: $P(\hat{p} \geq .75) \approx$ normcdf$(.75, \infty, .70, .0280) = .0371$



Conclude: About 3.71% of all SRSs of size 267 will give a sample proportion $\geq .75$.