Sampling Distribution of a Proportion

AP Statistics | Unit 3 | centre, spread and shape of $\hat{p}$

A population has a true proportion p. A random sample of size n gives a sample proportion $\hat{p}$, which varies from sample to sample. Its sampling distribution has three features.

Centre, spread, shape

Centre: $\mu_{\hat{p}} = p$ ($\hat{p}$ is an unbiased estimator of p).

Spread: $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$, provided $n \le 10\%$ of the population (the **10% condition**).

Shape: approximately normal if $np \ge 10$ and $n(1-p) \ge 10$ (the **Large Counts condition**).

1. The spread and the sample size

For $p = 0.60$: with $n = 100$, $\sigma_{\hat{p}} = \sqrt{\dfrac{(0.6)(0.4)}{100}} \approx 0.049$; with $n = 400$, $\sigma_{\hat{p}} \approx 0.0245$. To cut the standard deviation in half you need **four times** the sample, because n is under a square root.

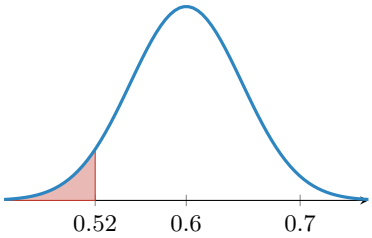
2. Why Large Counts matters

With $p = 0.6$, $n = 100$: $np = 60$ and $n(1-p) = 40$, both at least 10, so $\hat{p}$ is approximately normal. With $p = 0.05$, $n = 40$: $np = 2$; the distribution is piled against 0 and skewed right, and a normal model would be wrong.

3. Finding a probability

60% of voters support a candidate. In a random sample of 100 voters, how likely is a sample proportion of 0.52 or less?

1. **State** the distribution: $\hat{p}$ is approximately normal with $\mu_{\hat{p}} = 0.60$ and $\sigma_{\hat{p}} \approx 0.049$.
2. **Check**: $np = 60 \ge 10$ and $n(1-p) = 40 \ge 10$; 100 is less than 10% of all voters.
3. **Standardise**: $z = \dfrac{0.52 - 0.60}{0.049} \approx -1.63$.
4. **Answer**: $P(\hat{p} \le 0.52) \approx 0.051$.



About 5% of honest polls of this size miss by that much: unusual, not impossible.

4. A difference of two proportions

For independent random samples from two populations:

$$\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2, \qquad \sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}.$$

© 2026 S&D Academy

The *variances add*, even though it is a difference. It is approximately normal when Large Counts holds in both samples. Example: $p_1 = 0.60$, $p_2 = 0.50$, $n_1 = n_2 = 100$ gives mean 0.10 and standard deviation ≈ 0.070.

5. Closer with bigger n, and the 10% condition

With $p = 0.6$, the chance that $\hat{p}$ lands within 0.05 of p is about 0.69 for $n = 100$ ($\sigma_{\hat{p}} \approx 0.049$) and about 0.96 for $n = 400$ ($\sigma_{\hat{p}} \approx 0.0245$): a larger sample keeps the same centre but makes $\hat{p}$ far more likely to land close to p.

$\sqrt{p(1-p)/n}$ assumes independent draws. When the sample is more than 10% of the population (150 of 600 students), sampling without replacement makes the real spread smaller, so the formula *overstates* the variability; surveying everyone would give $\hat{p} = p$ exactly. When comparing two proportions, check Large Counts in each sample: $p = 0.05$, $n = 150$ gives $np = 7.5 < 10$, and one failing sample is enough to rule out the normal model.

Where students lose points

1. Using a normal model without checking Large Counts (with the actual numbers).
2. Forgetting the 10% condition.
3. Mixing up p and $\hat{p}$: p goes in the formula; $\hat{p}$ is the value asked about.
4. "Double the sample to halve the spread" (it takes four times).
5. Subtracting variances for a difference.

Now practise it

Theory sticks once you use it. On **sad-academy.com** you'll find practice questions for this exact subtopic in the AP format, and an AI that grades your free-response answers against the scoring guidelines, the way an AP reader would.

sad-academy.com

Data are hypothetical. AP® is a registered trademark of College Board, which is not affiliated with and does not endorse this material.

 © 2026 S&D Academy