Sampling Distribution of $\hat{p}$

AP Statistics · Mr. Merrick · January 12, 2026

Suppose we take a random sample of size n from a population.

- Let X be the number of successes in the sample.
- Let p be the true population proportion of successes.
- For a finite population of size N with K successes, $p = K/N$.
- The sample proportion is

$$\hat{p} = \frac{X}{n}.$$

Because the sample is random, X (and therefore $\hat{p}$) is a **random variable**. To make probability statements about $\hat{p}$, we must choose an appropriate probability model based on:

- whether observations can be treated as **independent**, and
- whether the **success–failure (normality) condition** is satisfied.

Row meaning: whether the **normality (success–failure) condition** is satisfied.
Column meaning: whether the **independence condition** is satisfied.

	Independence OK *(with replacement or $n \le 0.10N$)*	**Independence NOT OK** *(without replacement and $n > 0.10N$)*
Normality NOT OK *(success–failure fails)*	**Binomial (exact)** $X \sim \text{Bin}(n, p)$ $\hat{p} = X/n$ Use exact binomial probabilities.	**Hypergeometric (exact)** $X \sim \text{Hypergeometric}(N, K, n)$ $\hat{p} = X/n$ Use exact hypergeometric probabilities.
Normality OK *(success–failure holds)*	**Normal model for $\hat{p}$ (approx)** If independence is reasonable and $np \ge 10$, $n(1-p) \ge 10$, then $$\hat{p} \approx \mathcal{N}\left(p,\ \sqrt{\frac{p(1-p)}{n}}\right).$$ Use z-scores.	**Normal approximation (finite population)** When success–failure is met but the 10% condition fails, $$\hat{p} \approx \mathcal{N}\left(p,\ \sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}\right)$$ $$p = \frac{K}{N}.$$

1. For each situation below, identify the **most appropriate probability model** for the sampling distribution of $\hat{p}$. *No probability calculations are required. Justify your choice.*

   (a) A public-opinion researcher randomly surveys $n = 10$ adults from a very large country. Historically, about 4% of adults support a particular third-party candidate.

   Solution. The population is very large, so independence is reasonable. However, the success–failure condition fails since $np = 10(0.04) = 0.4 < 10$. Therefore, the normal model is not appropriate. Use the **binomial distribution (exact)** for X, and thus for $\hat{p}$.

   (b) A national polling organization surveys $n = 500$ randomly selected voters to estimate the proportion who approve of a proposed federal law. The true approval rate is 51%.

   Solution. The sample is random and 500 is far less than 10% of the population, so independence is reasonable. The success–failure condition is met since $np = 500(0.51) = 255$ and $n(1-p) = 500(0.49) = 245$. Therefore, the **normal model for $\hat{p}$** is appropriate.

   (c) A teacher randomly selects $n = 18$ students *without replacement* from a class of $N = 30$ students to check who completed a summer assignment. Exactly $K = 12$ students in the class completed the assignment.

   Solution. Sampling is without replacement and $18 > 0.10(30)$, so the independence condition is not satisfied. Therefore, neither the binomial nor the standard normal model is appropriate. Use the **hypergeometric distribution (exact)** for X and hence for $\hat{p}$.

   (d) A basketball player practices by shooting $n = 7$ free throws. Based on past data, the probability of making a free throw is 0.85, and shots are assumed independent.

   Solution. Shots are assumed independent. However, the success–failure condition fails since $n(1-p) = 7(0.15) = 1.05 < 10$. Thus, the normal model is not appropriate. Use the **binomial distribution (exact)**.

   (e) A shipment of $N = 60$ electronic components contains $K = 8$ defective items. A quality-control inspector randomly selects $n = 12$ components without replacement.

   Solution. Sampling is without replacement and $12 > 0.10(60)$, so independence is not satisfied. Therefore, the appropriate model is the **hypergeometric distribution (exact)**.

   (f) A town has $N = 500$ residents. A random sample of $n = 80$ residents is selected to estimate the proportion who own at least one dog. The true proportion is 0.40.

   Solution. Sampling is without replacement and $80 > 0.10(500)$, so the independence condition fails. Even though the success–failure condition is met, the appropriate model is the **hypergeometric distribution**.

2. A school counselor believes that 65% of students participate in at least one club. A random sample of $n = 120$ students is selected from a school with 1,800 students.

   (a) Identify the appropriate probability model for the sampling distribution of $\hat{p}$.

   Solution. The problem states that a random sample is taken. Since $120 \leq 0.10(1800)$, the independence condition is satisfied. The success–failure condition holds because $np = 120(0.65) = 78 \geq 10$ and $n(1-p) = 120(0.35) = 42 \geq 10$. Therefore, the **normal model for** $\hat{p}$ is appropriate.

   (b) Find the probability that the sample proportion of students who participate in at least one club is *greater than* 0.70.

   Solution. Using the normal model,

$$\hat{p} \approx \mathcal{N}\left(0.65,\ \sqrt{\frac{0.65(0.35)}{120}}\right).$$

   The standard deviation is

$$\sqrt{\frac{0.65(0.35)}{120}} \approx 0.0435.$$

   The z-score is

$$z = \frac{0.70 - 0.65}{0.0435} \approx 1.15.$$

   Thus,

$$P(\hat{p} > 0.70) \approx P(Z > 1.15) \approx 0.125.$$

3. A box contains $N = 90$ lightbulbs, $K = 15$ of which are defective. A technician randomly inspects $n = 20$ bulbs without replacement.

   (a) Explain why the normal model for $\hat{p}$ is *not* appropriate.

   Solution. The bulbs are sampled without replacement. Since $20 > 0.10(90)$, the independence condition is not satisfied. Because independence fails, the normal model for $\hat{p}$ is not appropriate.

   (b) Find the probability that *at most 3* of the inspected bulbs are defective.

   Solution. Let X be the number of defective bulbs in the sample. Then

$$X \sim \text{Hypergeometric}(90, 15, 20).$$

$$P(X \leq 3) = \sum_{x=0}^{3} \frac{\binom{15}{x}\binom{75}{20-x}}{\binom{90}{20}}.$$

4. A news organization claims that 48% of adults in a certain state support a proposed ballot initiative. The organization randomly surveys $n = 400$ adults from the state.

(a) Explain why the normal model for $\hat{p}$ is appropriate.

Solution. The sample is random and 400 is less than 10% of the adult population, so independence holds. The success–failure condition holds since $np = 400(0.48) = 192$ and $n(1-p) = 400(0.52) = 208$. Therefore, the normal model for $\hat{p}$ is appropriate.

(b) Find the probability that the sample proportion exceeds 0.52.

Solution.

$$\hat{p} \approx \mathcal{N}\left(0.48,\ \sqrt{\frac{0.48(0.52)}{400}}\right), \quad \text{SD} \approx 0.0250.$$

$$z = \frac{0.52 - 0.48}{0.0250} \approx 1.60, \quad P(\hat{p} > 0.52) \approx 0.055.$$

(c) Interpret the result.

Solution. A sample proportion of 0.52 or higher would occur about 5.5% of the time if the claim were true. This provides some evidence against the claim, but not strong evidence at the 5% level.

5. A manufacturing company claims that 30% of the items it produces fail an initial quality inspection. A random sample of $n = 150$ items is selected. Determine the probability that the sample proportion of failures is *less than* 0.25. Justify your model choice.

Solution. Model justification:

The problem states that a random sample of items is selected, so the *random condition* is satisfied. Because the sample size $n = 150$ is less than 10% of the total number of items produced, the *independence condition* is satisfied. The *success–failure condition* is met since

$$np = 150(0.30) = 45 \geq 10 \quad \text{and} \quad n(1-p) = 150(0.70) = 105 \geq 10.$$

Therefore, the sampling distribution of $\hat{p}$ can be approximated by a normal distribution.

Probability calculation:

Using the normal model,

$$\hat{p} \approx \mathcal{N}\left(0.30,\ \sqrt{\frac{0.30(0.70)}{150}}\right), \quad \text{SD} = \sqrt{\frac{0.21}{150}} \approx 0.0374.$$

$$z = \frac{0.25 - 0.30}{0.0374} \approx -1.34.$$

Thus,

$$P(\hat{p} < 0.25) \approx P(Z < -1.34) \approx 0.090.$$