8 Tests of Hypotheses Based on a Single Sample

8.3 Tests Concerning a Population Proportion

Tests Concerning a Population Proportion

Let p denote the proportion of individuals or objects in a population who possess a specified property (e.g., cars with manual transmissions or smokers who smoke a filter cigarette).

If an individual or object with the property is labeled a success (S), then p is the population proportion of successes.

Tests concerning p will be based on a random sample of size n from the population. Provided that n is small relative to the population size, X (the number of S's in the sample) has (approximately) a binomial distribution.

Tests Concerning a Population Proportion

Furthermore, if n itself is large [$np \geq 10$ and $n(1 - p) \geq 10$], both X and the estimator $\hat{p} = X/n$ are approximately normally distributed.

We first consider large-sample tests based on this latter fact and then turn to the small sample case that directly uses the binomial distribution.

Large-Sample Tests

Large-Sample Tests

Large-sample tests concerning p are a special case of the more general large-sample procedures for a parameter θ.

Let $\hat{\theta}$ be an estimator of θ that is (at least approximately) unbiased and has approximately a normal distribution.

The null hypothesis has the form H_0: $\theta = \theta_0$ where θ_0 denotes a number (the null value) appropriate to the problem context.

Large-Sample Tests

Suppose that when H_0 is true, the standard deviation of $\hat{\theta}$, $\sigma_{\hat{\theta}}$, involves no unknown parameters.

For example, if $\theta = \mu$ and $\hat{\theta} = \overline{X}$, $\sigma_{\hat{\theta}} = \sigma_{\overline{X}} = \sigma/\sqrt{n}$, which involves no unknown parameters only if the value of σ is known.

Large-Sample Tests

A large-sample test statistic results from standardizing $\hat{\theta}$ under the assumption that H_0 is true (so that $E(\hat{\theta}) = \theta_0$):

$$\text{Test statistic: } Z = \frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}}$$

If the alternative hypothesis is H_a: $\theta > \theta_0$, an upper-tailed test whose significance level is approximately α is specified by the rejection region $z \geq z_\alpha$.

The other two alternatives, H_a: $\theta < \theta_0$ and H_a: $\theta \neq \theta_0$, are tested using a lower-tailed z test and a two-tailed z test, respectively.

Large-Sample Tests

In the case $\theta = p$, $\sigma_{\hat{\theta}}$ will not involve any unknown parameters when H_0 is true, but this is atypical.

When $\sigma_{\hat{\theta}}$ does involve unknown parameters, it is often possible to use an estimated standard deviation $S_{\hat{\theta}}$ in place of $\sigma_{\hat{\theta}}$ and still have *Z* approximately normally distributed when H_0 is true (because when *n* is large, $s_{\hat{\theta}} \approx \sigma_{\hat{\theta}}$ for most samples).

The large-sample test we have seen earlier furnishes an example of this: Because σ is usually unknown, we use $s_{\hat{\theta}} = s_{\overline{X}} = s/\sqrt{n}$ in place of $\sigma/\sqrt{n}$ in the denominator of *z*.

Large-Sample Tests

The estimator $\hat{p} = X/n$ is unbiased $(E(\hat{p}) = p)$, has approximately a normal distribution, and its standard deviation is $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$.

x/n

When H_0 is true, $E(\hat{p}) = p_0$ and $\sigma_{\hat{p}} = \sqrt{p_0(1-p_0)/n}$, so $\sigma_{\hat{p}}$ does not involve any unknown parameters. It then follows that when n is large and H_0 is true, the test statistic

$$\boxed{Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}}$$

ค่าที่เคลม

☆ Sample

has approximately a standard normal distribution.

Large-Sample Tests

If the alternative hypothesis is H_a: $p > p_0$ and the upper-tailed rejection region $z \geq z_\alpha$ is used, then

$$P(\text{type I error}) = P(H_0 \text{ is rejected when it is true})$$

$$= P(Z \geq z_\alpha \text{ when } Z \text{ has approximately a standard normal distribution}) \approx \alpha$$

Thus the desired level of significance α is attained by using the critical value that captures area α in the upper tail of the z curve.

Large-Sample Tests

Rejection regions for the other two alternative hypotheses, lower-tailed for H_a: $p < p_0$ and two-tailed for H_a: $p \neq p_0$, are justified in an analogous manner.

Null hypothesis: H_0: $p = p_0$

Test statistic value: $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$

Large-Sample Tests

Alternative Hypothesis	**Rejection Region**
H_a: $p > p_0$	$z \geq z_\alpha$ (upper-tailed)
H_a: $p < p_0$	$z \leq -z_\alpha$ (lower-tailed)
H_a: $p \neq p_0$	either $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$ (two-tailed)

These test procedures are valid provided that $np_0 \geq 10$ and $n(1 - p_0) \geq 10$.

Example 11

Natural cork in wine bottles is subject to deterioration, and as a result wine in such bottles may experience contamination.

The article "Effects of Bottle Closure Type on Consumer Perceptions of Wine Quality" (*Amer*. *J*. *of Enology and Viticulture*, 2007: 182–191) reported that, in a tasting of commercial chardonnays, 16 of 91 bottles were considered spoiled to some extent by cork-associated characteristics.

Does this data provide strong evidence for concluding that more than 15% of all such bottles are contaminated in this way?

Example 11

cont'd

Let's carry out a test of hypotheses using a significance level of .10. $= \alpha$

1. p = the true proportion of all commercial chardonnay bottles considered spoiled to some extent by cork-associated characteristics.

2. The null hypothesis is H_0: $p = .15$.

3. The alternative hypothesis is H_a: $p > .15$, the assertion that the population percentage exceeds 15%. $\alpha = ?$

Example 11

cont'd

4. Since $np_0 = 91(.15) = 13.65 > 10$ and $nq_0 = 91(.85) = 77.35 > 10$, the large-sample z test can be used. The test statistic value is

$$z = (\hat{p} - .15)/\sqrt{(.15)(.85)/n}.$$

p_0 q_0

5. The form of H_a implies that an upper-tailed test is appropriate: Reject H_0 if $z \geq z_{.10} = 1.28$.

6. $\hat{p} = 16/91 = .1758$, from which

$$z = (.1758 - .15)/\sqrt{(.15)(.85)/91} = .0258/.0374 = .69$$

p_0

Example 11

cont'd

7. Since .69 < 1.28, z is not in the rejection region. At significance level .10, the null hypothesis cannot be rejected.

Although the percentage of contaminated bottles in the sample somewhat exceeds 15%, the sample percentage

is not large enough to conclude that the population percentage exceeds 15%.

The difference between the sample proportion .1758 and the null value .15 can adequately be explained by sampling variability.

Large-Sample Tests

***β* and Sample Size Determination** When H_0 is true, the test statistic Z has approximately a standard normal distribution.

Now suppose that H_0 is *not* true and that $p = p'$. Then Z still has approximately a normal distribution (because it is a linear function of $\hat{p}$), but its mean value and variance are no longer 0 and 1, respectively. Instead,

$$E(Z) = \frac{p' - p_0}{\sqrt{p_0(1 - p_0)/n}} \qquad V(Z) = \frac{p'(1 - p')/n}{p_0(1 - p_0)/n}$$

The probability of a type II error for an upper-tailed test is $\beta(p') = P(Z > z_\alpha$ when $p = p')$.

Large-Sample Tests

This can be computed by using the given mean and variance to standardize and then referring to the standard normal cdf.

In addition, if it is desired that the level α test also have $\beta(p')=\beta$ for a specified value of β, this equation can be solved for the necessary n.

Large-Sample Tests

General expressions for $\beta(p')$ and n are given in the accompanying box.

Alternative Hypothesis	$\boldsymbol{\beta(p')}$
H_a: $p > p_0$	$\Phi\left[\dfrac{p_0 - p' + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right]$
H_a: $p < p_0$	$1 - \Phi\left[\dfrac{p_0 - p' - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right]$
H_a: $p \neq p_0$	$\Phi\left[\dfrac{p_0 - p' + z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right]$ $-\Phi\left[\dfrac{p_0 - p' - z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right]$

Large-Sample Tests

The sample size n for which the level α test also satisfies $\beta(p') = \beta$ is

$$n = \begin{cases} \left[\dfrac{z_\alpha\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p'(1-p')}}{p'-p_0}\right]^2 & \text{one-tailed test} \\[2ex] \left[\dfrac{z_{\alpha/2}\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p'(1-p')}}{p'-p_0}\right]^2 & \text{two-tailed test (an approximate solution)} \end{cases}$$

Example 12

A package-delivery service advertises that at least 90% of all packages brought to its office by 9 A.M. for delivery in the same city are delivered by noon that day.

Let p denote the true proportion of such packages that are delivered as advertised and consider the hypotheses H_0: $p = .9$ versus H_a: $p < .9$.

If only 80% of the packages are delivered as advertised, how likely is it that a level .01 test based on $n = 225$ packages will detect such a departure from H_0? What should the sample size be to ensure that $\beta(.8) = .01$?

Example 12

cont'd

With $\alpha = .01$, $p_0 = .9$, $p' = .8$, and $n = 225$,

$$\beta(.8) = 1 - \Phi\left(\frac{.9 - .8 - 2.33\sqrt{(.9)(.1)/225}}{\sqrt{(.8)(.2)/225}}\right)$$

$$= 1 - \Phi(2.00) = .0228$$

Thus the probability that H_0 will be rejected using the test when $p = .8$ is .9772— roughly 98% of all samples will result in correct rejection of H_0.

Example 12

cont'd

Using $z_\alpha = z_\beta = 2.33$ in the sample size formula yields

$$n = \left[\frac{2.33\sqrt{(.9)(.1)} + 2.33\sqrt{(.8)(.2)}}{.8 - .9} \right]^2 \approx 266$$