

Medical Statistics and Data Analysis

Lecture 8: Hypothesis Testing: One-Sample Inference

Zhiyu Wan, PhD (wanzhy@shanghaitech.edu.cn)

Assistant Professor of Biomedical Engineering

ShanghaiTech University

October 21, 2026

Learning Goals of This Lecture

- One-Sample Inference
- p -value
- Statistical Power Estimation
- Sample Size Estimation

Introduction

- Hypothesis-testing framework specifies two hypotheses:
- **Null and alternative hypothesis**
- Hypothesis-testing provides an objective framework for making decisions using probabilities methods, rather than relying on subjective impressions.
- It provides a uniform decision-making criterion that is consistent.
- In a **one-sample problem**, hypotheses are specified about a single distribution.
- In a **two-sample problem**, two different distributions are compared.

- The **null hypothesis**, denoted by H_0 , is the hypothesis that is to be tested. The **alternative hypothesis**, denoted by H_1 , is the hypothesis that in some sense contradicts the null hypothesis.

$$H_0: \mu = \mu_0 \quad \text{vs.} \quad H_1: \mu < \mu_0$$

- Suppose the only possible decisions are whether H_0 is true or H_1 is true. For ease of notation, all outcomes in a hypothesis testing situation generally refer to the null hypothesis: If we decide H_0 is true, then we say we **accept H_0** (or fail to reject H_0). If we decide H_1 is true, then we state that H_0 is not true or, equivalently, that we **reject H_0** .
- Thus, four possible outcomes can occur:

TABLE 7.1 Four possible outcomes in hypothesis testing

		Truth	
		H_0	H_1
Decision	Accept H_0	H_1 is true H_0 is accepted	Type II Error, β
	Reject H_0	Type I Error, α	H_1 is true H_0 is rejected

- **The probability of a type I error** is the probability of rejecting the null hypothesis when H_0 is true.
- It is denoted by α and is commonly referred to as the **significance level** of a test.
- **The probability of a type II error** is the probability of accepting the null hypothesis when H_1 is true.
- This probability is a function of μ as well as other factors. It is usually denoted by β .
- The **(statistical) power of a test** is defined as
$$1 - \beta = 1 - \text{probability of a type II error}$$
$$= P(\text{rejecting } H_0 \mid H_1 \text{ true})$$
- The aim in hypothesis testing is to use statistical tests that make α and β as small as possible, which means rejecting and accepting null hypothesis less often, respectively.

One-Sample Test for the Mean of a Normal Distribution: One-Sided Alternatives

- **Acceptance region:** range of values of x for which H_0 is accepted.
- **Rejection region:** range of values of x for which H_0 is rejected.
- If $\bar{x}$ is sufficiently smaller than μ_0 , then H_0 is rejected; otherwise, H_0 is accepted.
- If H_0 is true: the most likely values of $\bar{x}$ cluster around μ_0 .
- If H_1 is true: the most likely values of $\bar{x}$ cluster around μ_1 .
- Such a test based on sample mean has the highest power among all tests with a given type I error of α .
- A **one-tailed test** is a test in which the values of the parameter being studied (in this case μ) under the alternative hypothesis are allowed to be either greater than or less than the values of the parameter under the null hypothesis (μ_0), but not both.

One-Sample t Test for the Mean of a Normal Distribution with Unknown Variance (Alternative Mean $<$ Null Mean)

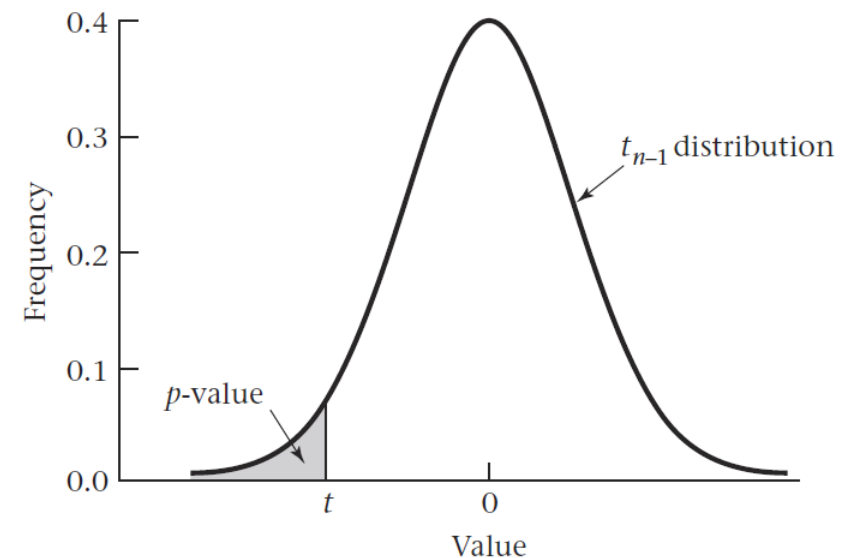
- To test the hypothesis $H_0: \mu = \mu_0$, σ unknown vs. $H_1: \mu < \mu_0$, σ unknown with a significance level of α , we compute

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

- If $t \leq t_{n-1,\alpha}$, then we **reject H_0** .
- If $t > t_{n-1,\alpha}$, then we **accept H_0** (or **fail to reject H_0**).
- The value of t is called a **test statistic** because the test procedure is based on this statistic.
- The value $t_{n-1,\alpha}$ is called a **critical value** because the outcome of the test depends on whether the test statistic $t \leq t_{n-1,\alpha}$ = critical value, whereby we reject H_0 , or $t > t_{n-1,\alpha}$, whereby we accept H_0 .

- The general approach in which we compute a test statistic and determine the outcome of a test by comparing the test statistic with a critical value determined by the type I error is called the **critical-value method** of hypothesis testing.
- The **p -value** for any hypothesis test is the **α level** at which we would be indifferent between accepting or rejecting H_0 given the sample data at hand. That is, the p -value is at the α level at which the given value of the test statistic (such as t) is on the borderline between the acceptance and rejection regions.
- Solving for p as a function of t
we get $p = \Pr(t_{n-1} \leq t)$

FIGURE 7.1 Graphic display of a p -value



- The ***p*-value** can also be thought of as the probability of obtaining a test statistic as extreme as or more extreme than the actual test statistic obtained, given that the null hypothesis is true:

$$p = \Pr(\text{observed or more extreme data} \mid H_0)$$

- The ***p*-value** indicates exactly how significant the results are without performing repeated significance tests at different α levels.

Guidance for Judging the Significance of a *p*-Value

- If $0.01 < p \leq 0.05$, then the results are **significant**.
- If $0.001 < p \leq 0.01$, then the results are highly significant.
- If $p \leq 0.001$, then the results are very highly significant.
- If $p > 0.05$, then the results are considered **not statistically significant** (denoted as **NS**).
- If $0.05 < p \leq 0.1$, then a trend toward statistical significance is sometimes noted.

Determination of Statistical Significance for Results from Hypothesis Tests

Either of the following methods can be used to establish whether results from hypothesis tests are statistically significant:

1. The test statistic t can be computed and compared with the critical value $t_{n-1,\alpha}$ at an α level of 0.05. If $H_0: \mu = \mu_0$ vs. $H_1: \mu < \mu_0$ is being tested and $t \leq t_{n-1,0.05}$, then **H_0 is rejected** and the results are declared **statistically significant** ($p \leq 0.05$). Otherwise, **H_0 is accepted** and the results are declared **not statistically significant** ($p > 0.05$). This approach is called the **critical-value method**.
2. The exact p -value can be computed and, if $p \leq 0.05$, then H_0 is rejected and the results are declared **statistically significant**. Otherwise, if $p > 0.05$, then H_0 is accepted and the results are declared **not statistically significant**. This approach is called the **p -value method**.

One-Sample t Test for the Mean of a Normal Distribution with Unknown Variance (Alternative Mean $>$ Null Mean)

- To test the hypothesis

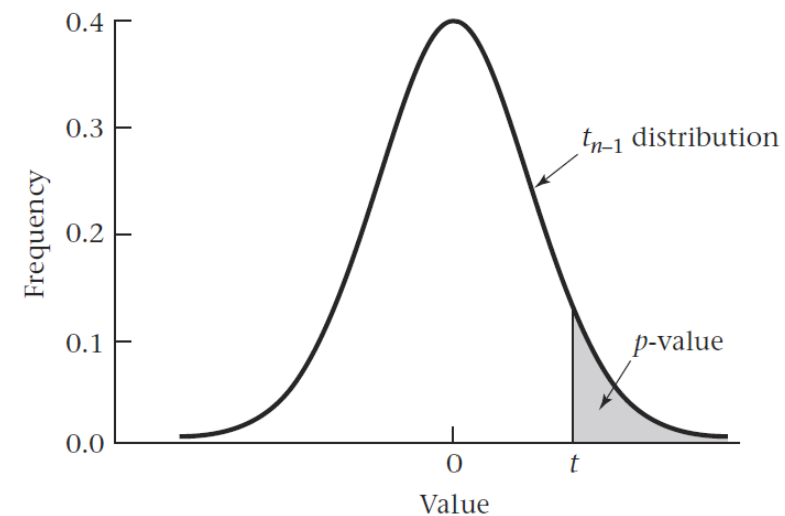
$$H_0: \mu = \mu_0 \quad \text{vs.} \quad H_1: \mu > \mu_0$$

- with a significance level of α , the best test is based on t , where

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- If $t \geq t_{n-1, 1-\alpha}$, then H_0 is rejected.
- If $t < t_{n-1, 1-\alpha}$, then H_0 is accepted.
- The p -value for this test is given by
 $p = \Pr(t_{n-1} > t)$

FIGURE 7.2 p -value for the one-sample t test when the alternative mean (μ_1) $>$ null mean (μ_0)



One-Sample Test for the Mean of a Normal Distribution: Two-Sided Alternatives

- A **two-tailed test** is a test in which the values of the parameter being studied (in this case μ) under the alternative hypothesis are allowed to be either greater than or less than the values of the parameter under the null hypothesis (μ_0).
- A reasonable decision rule to test for alternatives on either side of the null mean is to reject H_0 if it is either too small or too large. Another way of stating the rule is that H_0 will be rejected if t is either $\leq c_1$ or $\geq c_2$ for some constants c_1, c_2 , and H_0 will be accepted if $c_1 < t < c_2$.

$$\Pr(\text{reject } H_0 \mid H_0 \text{ true}) = \Pr(t \leq c_1 \text{ or } t \geq c_2 \mid H_0 \text{ true})$$

$$= \Pr(t \leq c_1 \mid H_0 \text{ true}) + \Pr(t \geq c_2 \mid H_0 \text{ true}) = \alpha$$

$$\Pr(t \leq c_1 \mid H_0 \text{ true}) = \Pr(t \geq c_2 \mid H_0 \text{ true}) = \frac{\alpha}{2}$$

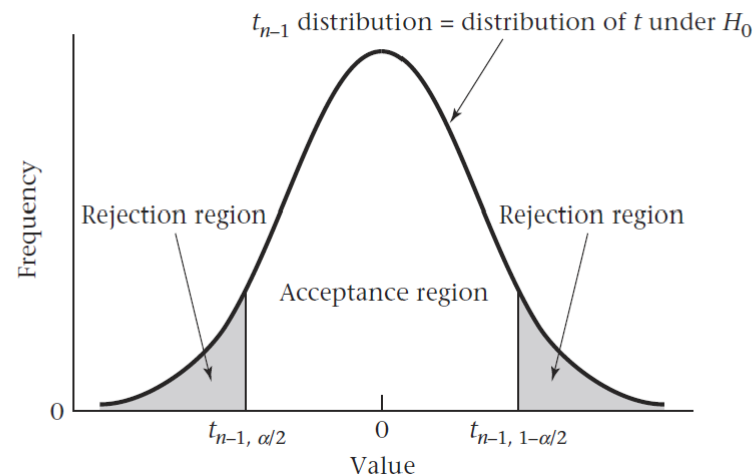
One-sample t Test for the Mean of a Normal Distribution with Unknown Variance (Two-Sided Alternative)

- To test the hypothesis $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ with a significance level of α , the best test is based on

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}.$$

- If $|t| \geq t_{n-1, 1-\alpha/2}$, then **H_0 is rejected**.
- If $|t| < t_{n-1, 1-\alpha/2}$, then **H_0 is accepted**.

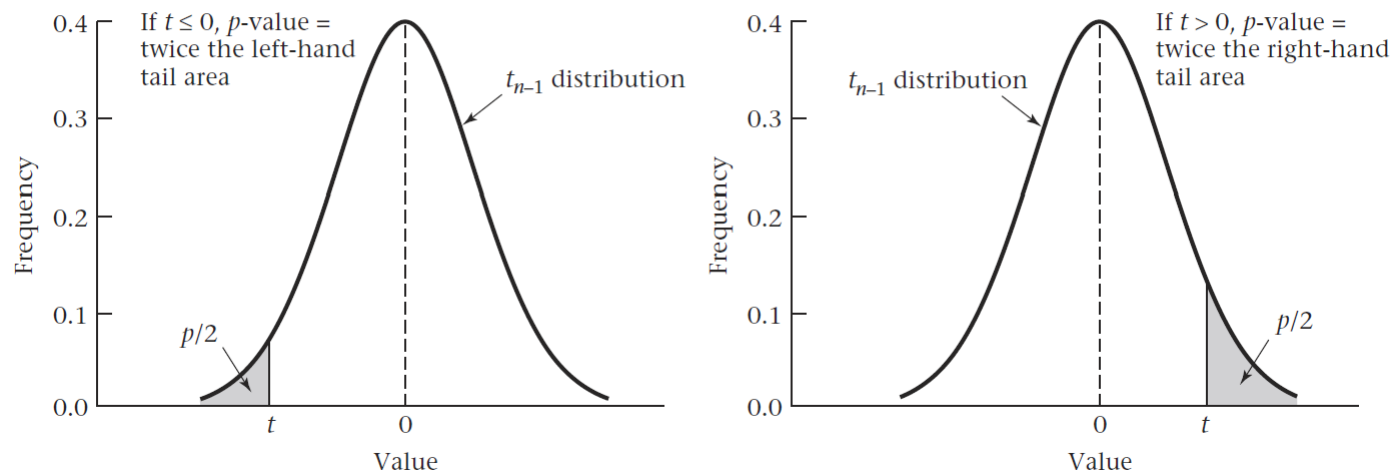
FIGURE 7.3 One-sample t test for the mean of a normal distribution (two-sided alternative)



p -value for the One-Sample t Test for the Mean of a Normal Distribution (Two-Sided Alternative)

- Let $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ and $p = 2 \times \Pr(t_{n-1} \geq |t|) = \begin{cases} 2 \times \Pr(t_{n-1} \leq t), & \text{if } t \leq 0 \\ 2 \times [1 - \Pr(t_{n-1} \leq t)], & \text{if } t > 0 \end{cases}$
- The **p -value** is the probability under the null hypothesis of obtaining a test statistic as extreme as or more extreme than the observed test statistic, where, because a **two-sided alternative hypothesis** is being used, extremeness is measured by the absolute value of the test statistic.

FIGURE 7.4 Illustration of the p -value for a one-sample t test for the mean of a normal distribution (two-sided alternative)



When is a one-sided test more appropriate than a two-sided test?

- Generally, the sample mean falls in the expected direction from μ_0 and it is easier to reject H_0 using a one-sided test than using a two-sided test.
- A two-sided test can be more **conservative** because it is not necessary to guess the appropriate side of the null hypothesis for the alternative hypothesis.
- In some cases, only alternatives on one side of the null mean are **of interest** or are **possible**, and a one-sided test is better than a two-sided test because it has more **(statistical) power** (since it is **easier to reject** H_0 based on a finite sample if H_1 is actually true).
- The decision whether to use a one-sided or two-sided test must be made **before** the data analysis (or before data collection) begins so as not to bias conclusions based on results of hypothesis testing.
- Do not change from a two-sided to a one-sided test after looking at the data.

One-Sample z Test for the Mean of a Normal Distribution with Known Variance (Two-Sided Alternative)

- To test the hypothesis $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ with a significance level of α , where the underlying standard deviation σ is known, the best test is based on

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

- If $z \leq z_{\alpha/2}$ or $z \geq z_{1-\alpha/2}$, then **H_0 is rejected**.
- If $z_{\alpha/2} < z < z_{1-\alpha/2}$, then **H_0 is accepted**.
- To compute a two-sided p-value, we have

$$p = 2[1 - \Phi(|z|)] = \begin{cases} 2\Phi(z), & z \leq 0 \\ 2[1 - \Phi(z)], & z > 0 \end{cases}$$

where $\Phi(z)$ is the **CDF (cumulative distribution function)** of standard normal distribution:

$$\Phi(z) = P(Z \leq z), \quad Z \sim N(0,1)$$

One-Sample z Test for the Mean of a Normal Distribution with Known Variance (One-Sided Alternative)

$(\mu_1 < \mu_0)$ Left-Sided Test

- To test the hypothesis $H_0: \mu = \mu_0$ vs. $H_1: \mu < \mu_0$ with a significance level of α , where the underlying standard deviation σ is known, the best test is based on

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

- If $z \leq z_\alpha$, then H_0 is rejected; if $z > z_\alpha$, then H_0 is accepted. The p -value is given by $p = \Phi(z)$.

$(\mu_1 > \mu_0)$ Right-Sided Test

- To test the hypothesis $H_0: \mu = \mu_0$ vs. $H_1: \mu > \mu_0$ with a significance level of α , where the underlying standard deviation σ is known, the best test is based on

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

- If $z \geq z_{1-\alpha}$, then H_0 is rejected; if $z < z_{1-\alpha}$, then H_0 is accepted. The p -value is given by $p = 1 - \Phi(z)$.

The Power of a Test

- It tells us how likely it is that a statistically significant difference will be detected based on a finite sample size n , if the alternative hypothesis is true, that is, if the true mean μ differs from the mean under the null hypothesis (μ_0).

$$\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \geq z_{1-\alpha}$$

- Multiplying both sides with $\sigma/\sqrt{n}$, and adding μ_0 , we get $\bar{x} \geq \mu_0 + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$ or $\bar{x} \leq \mu_0 + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$.

Illustration of power for the one-sample test for the mean of a normal distribution with known variance ($\mu_1 < \mu_0$)

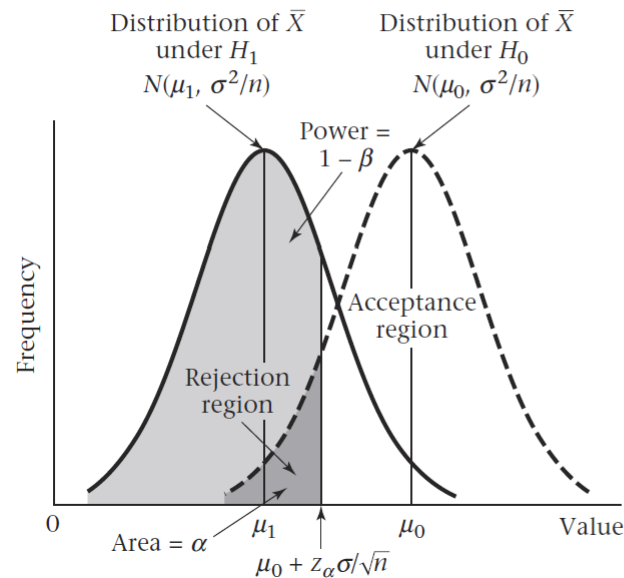
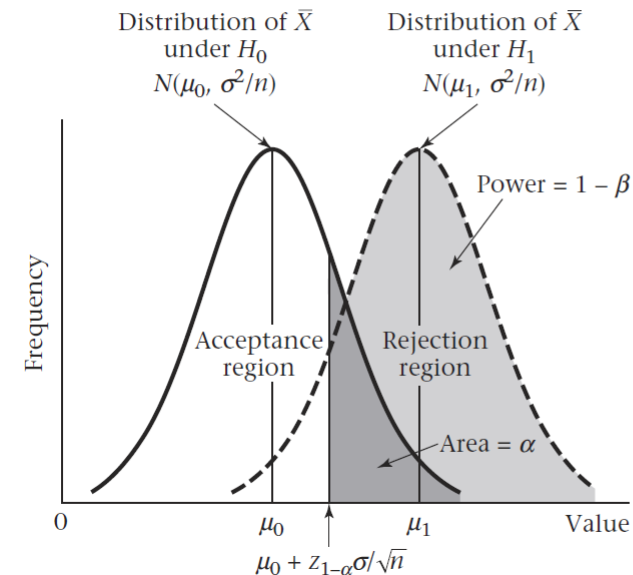


Illustration of power for the one-sample test for the mean of a normal distribution with known variance ($\mu_1 > \mu_0$)



Power for the One-Sample z Test for the Mean of a Normal Distribution with Known Variance (One-Sided Alternative)

- The power of the test for the hypothesis $H_0: \mu = \mu_0$ vs. $H_1: \mu = \mu_1$, where the underlying distribution is normal and the population variance (σ^2) is assumed known, is given by

$$\text{Power} = \Phi\left(z_\alpha + \frac{|\mu_0 - \mu_1|\sqrt{n}}{\sigma}\right) = \Phi\left(-z_{1-\alpha} + \frac{|\mu_0 - \mu_1|\sqrt{n}}{\sigma}\right)$$

Factors affecting the power

- If the significance level is made smaller (α decreases), z_α decreases and hence the power decreases.
- If the alternative mean is shifted farther away from the null mean ($|\mu_0 - \mu_1|$ increases), then the power increases.
- If the standard deviation of the distribution of individual observations increases (σ increases), then the power decreases.
- If the sample size increases (n increases), then the power increases.

Power for the One-Sample z Test for the Mean of a Normal Distribution (Two-Sided Alternative)

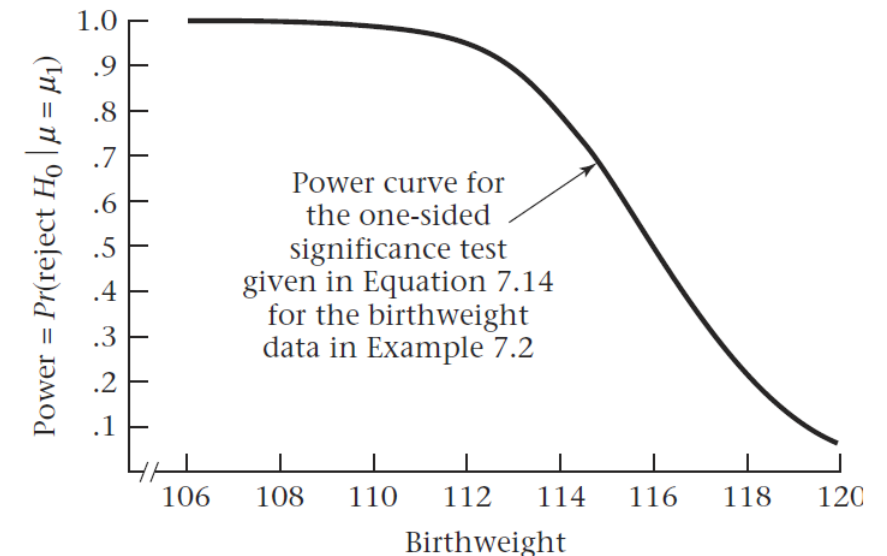
- The power of the two-sided test $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ for the specific alternative $\mu = \mu_1$, where the underlying distribution is normal and the population variance (σ^2) is assumed known, is given exactly by

$$\text{Power} = \Phi \left[-z_{1-\alpha/2} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma} \right] + \Phi \left[-z_{1-\alpha/2} + \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma} \right]$$

- and approximately by

$$\text{Power} \approx \Phi \left[-z_{1-\alpha/2} + \frac{|\mu_0 - \mu_1|\sqrt{n}}{\sigma} \right]$$

FIGURE 7.7 Power curve for the birthweight data in Example 7.2



- We can re-express the rejection criteria in terms of $\bar{x}$ as

$$\bar{x} \leq \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \text{or} \quad \bar{x} \geq \mu_0 + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

- The power of the test vs. the specific alternative $\mu = \mu_1$ is given by

$$\begin{aligned} \text{Power} &= P\left(\bar{X} \leq \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \mid \mu = \mu_1\right) + P\left(\bar{X} \geq \mu_0 + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \mid \mu = \mu_1\right) \\ &= \Phi\left(\frac{\mu_0 + z_{\alpha/2} \sigma / \sqrt{n} - \mu_1}{\sigma / \sqrt{n}}\right) + 1 - \Phi\left(\frac{\mu_0 + z_{1-\alpha/2} \sigma / \sqrt{n} - \mu_1}{\sigma / \sqrt{n}}\right) \\ &= \Phi\left[z_{\alpha/2} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma}\right] + 1 - \Phi\left[z_{1-\alpha/2} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma}\right] \end{aligned}$$

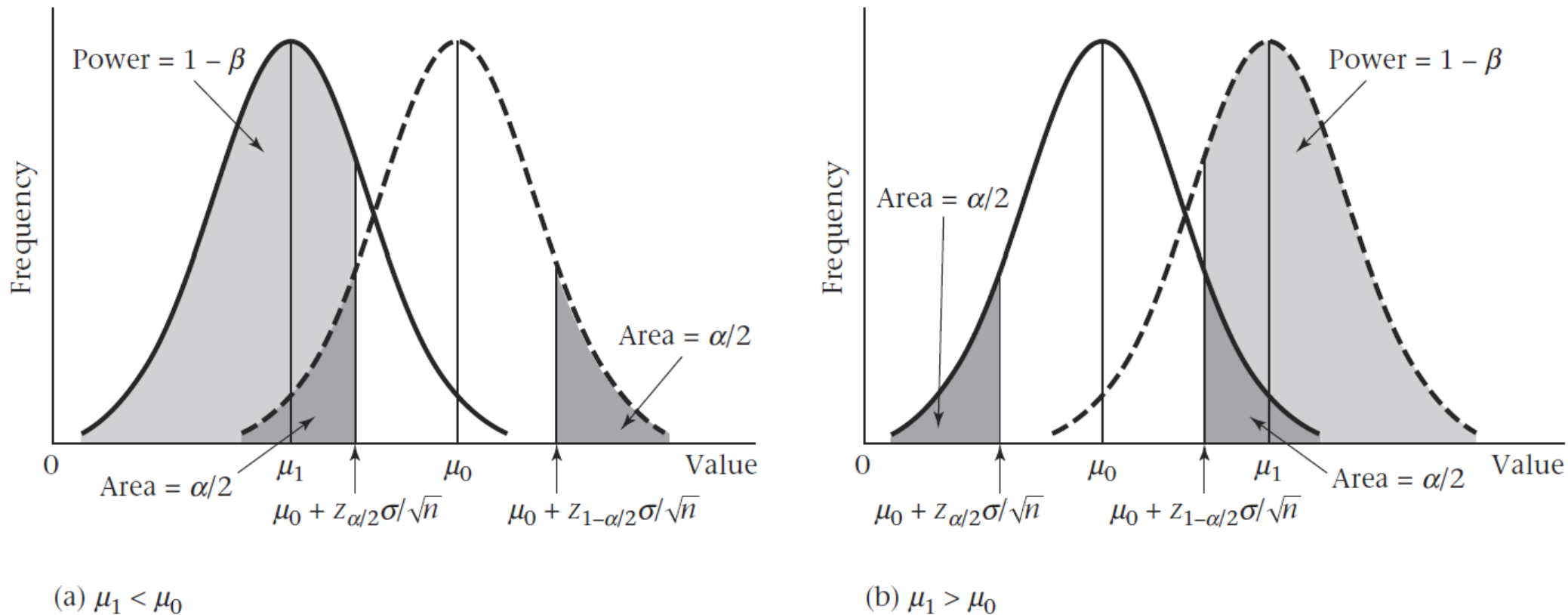
- Using $1 - \Phi(x) = \Phi(-x)$, the last two terms can be combined as follows:

$$\text{Power} = \Phi\left[z_{\alpha/2} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma}\right] + \Phi\left[-z_{1-\alpha/2} + \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right]$$

- Finally, recalling the relationship $z_{\alpha/2} = -z_{1-\alpha/2}$, we have

$$\text{Power} = \Phi\left[-z_{1-\alpha/2} + \frac{(\mu_0 - \mu_1)\sqrt{n}}{\sigma}\right] + \Phi\left[-z_{1-\alpha/2} + \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right]$$

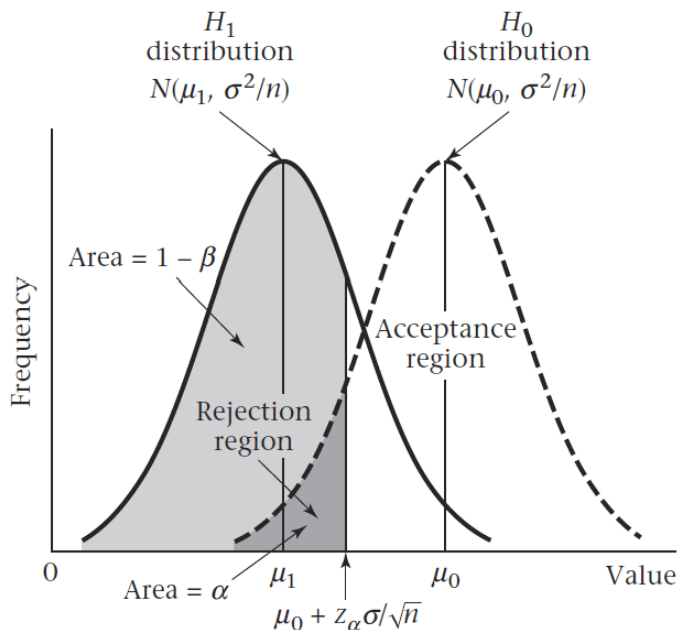
FIGURE 7.8 Illustration of power for a two-sided test for the mean of a normal distribution with known variance



Sample-Size Determination: One-Sided Alternatives

- Suppose we wish to test $H_0: \mu = \mu_0$ vs. $H_1: \mu = \mu_1$ where the data are normally distributed with mean μ and known variance σ^2 . The sample size n needed to conduct a one-sided test with significance level α and probability of detecting a significant difference $= 1 - \beta$

FIGURE 7.9 Requirements for appropriate sample size



$$n = \frac{\sigma^2 (z_{1-\beta} + z_{1-\alpha})^2}{(\mu_0 - \mu_1)^2}$$

Factors Affecting the Sample Size

$$n = \frac{\sigma^2(z_{1-\beta} + z_{1-\alpha})^2}{(\mu_0 - \mu_1)^2}$$

- The sample size increases as σ^2 increases.
- The sample size increases as the significance level is made smaller (α decreases).
- The sample size increases as the required power increases ($1 - \beta$ increases).
- The sample size decreases as the absolute value of the distance between the null and alternative means ($|\mu_0 - \mu_1|$) increases.

Sample-Size Estimation When Testing for the Mean of a Normal Distribution (Two-Sided Alternative)

- Suppose we wish to test $H_0: \mu = \mu_0$ vs. $H_1: \mu = \mu_1$ where the data are normally distributed with mean μ and known variance σ^2 . The sample size needed to conduct a two-sided test with significance level α and power $1 - \beta$ is

$$n = \frac{\sigma^2 (z_{1-\beta} + z_{1-\alpha/2})^2}{(\mu_0 - \mu_1)^2}$$

Sample-Size Estimation Based on CI Width

- Suppose we wish to estimate the mean of a normal distribution with sample variance s^2 and require that the two-sided $100\% \times (1 - \alpha)$ CI for μ be no wider than L . The number of subjects needed is approximately

$$n = \frac{4z_{1-\alpha/2}^2 s^2}{L^2}$$

The Relationship Between Hypothesis Testing and Confidence Intervals (Two-Sided Case)

- Suppose we are testing $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$. H_0 is rejected with a two-sided level α test if and only if the two-sided $100\% \times (1 - \alpha)$ CI for μ does not contain μ_0 . H_0 is accepted with a two-sided level α test if and only if the two-sided $100\% \times (1 - \alpha)$ CI for μ does contain μ_0 .
- The two-sided $100\% \times (1 - \alpha)$ CI for μ contains all values μ_0 such that we accept H_0 using a two-sided test with significance level α , where the hypotheses are $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$. Conversely, the $100\% \times (1 - \alpha)$ CI does not contain any value μ_0 for which we can reject H_0 using a two-sided test with significance level α , where $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$.

Bayesian Inference

$$\Pr\left(\bar{x} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} | \mathbf{x}\right) = (1 - \alpha)$$

- To use this, we need to specify a **prior distribution** for μ . A prior distribution is a generalization of the concept of a prior probability, where a probability distribution of possible values for a parameter (such as μ) is specified before looking at the available sample data.
- This distribution is then modified after some sample data are acquired to obtain a posterior distribution for μ .
- In the absence of data, we often have no knowledge of μ , so all possible values of μ are equally likely. This is called **flat** or **noninformative prior distribution** and is written as $Pr(\mu) \propto c$, where c is a constant.
- If distribution is normal and σ is known, then all the information in the sample concerning μ is contained in $\bar{x}$, that is, $Pr(\mu|x_1, \dots, x_n) = Pr(\mu|\bar{x})$.

- $\bar{X}$ is referred to as a **sufficient statistic** for μ . To obtain the posterior distribution, we use Bayes' rule.

$$\Pr(\mu | \bar{x}) = \frac{\Pr(\bar{x} | \mu) \Pr(\mu)}{\Pr(\bar{x})}$$

- However, the prior probability of μ , $\Pr(\mu)$, is the same for all values of μ .
- Thus,

$$\Pr(\mu | \bar{x}) \propto \Pr(\bar{x} | \mu)$$

- For a continuous distribution, we approximate $\Pr(\bar{x} | \mu)$ by using the pdf function $f(\bar{x} | \mu)$.

$$\begin{aligned} f(\bar{x} | \mu) &= \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}} \exp \left\{ -\frac{12^2}{\left[\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \right]^2} \right\} \\ &= \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}} \exp \left\{ -\frac{12^2}{\left[\frac{\mu - \bar{x}}{\sigma/\sqrt{n}} \right]^2} \right\} \end{aligned}$$

- It follows that

$$\Pr(\mu | \bar{x}) \propto \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}} \exp \left\{ -\frac{12^2}{\left[\frac{\mu - \bar{x}}{\sigma/\sqrt{n}} \right]^2} \right\}$$

- The distribution of μ given $\bar{x}$ is normally distributed with mean = $\bar{x}$ and variance σ^2/n , or

$$\mu | \bar{x} \sim N(\bar{x}, \sigma^2/n)$$

- A $100\% \times (1 - \alpha)$ **Bayesian credible interval** for μ based on the sample data of the form $Pr(\mu_1 < \mu < \mu_2 | \mathbf{X}) = (1 - \alpha)$ where

$$\mu_1 = \bar{X} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \mu_2 = \bar{X} + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

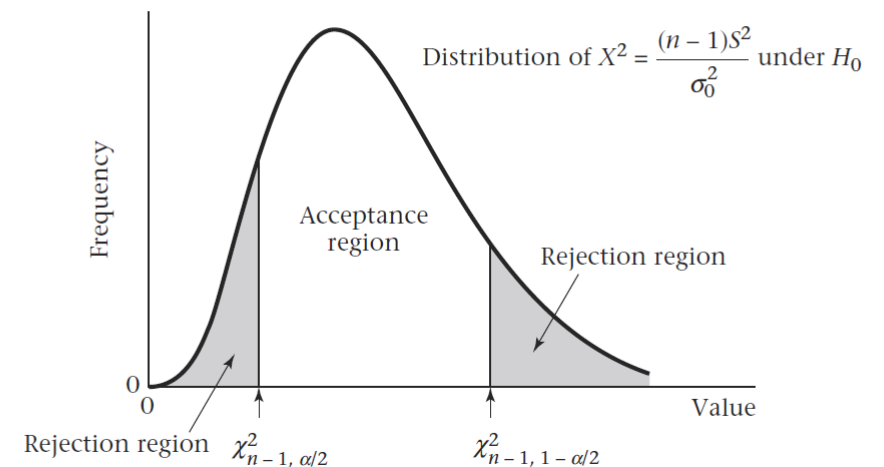
One-Sample χ^2 Test for the Variance of a Normal Distribution (Two-Sided Alternative)

- We compute the test statistic

$$X^2 = \frac{(n - 1)s^2}{\sigma_0^2}$$

- If $X^2 < \chi_{n-1, \alpha/2}^2$ or $X^2 > \chi_{n-1, 1-\alpha/2}^2$ then H_0 is rejected.
- If $\chi_{n-1, \alpha/2}^2 \leq X^2 \leq \chi_{n-1, 1-\alpha/2}^2$, then H_0 is accepted.

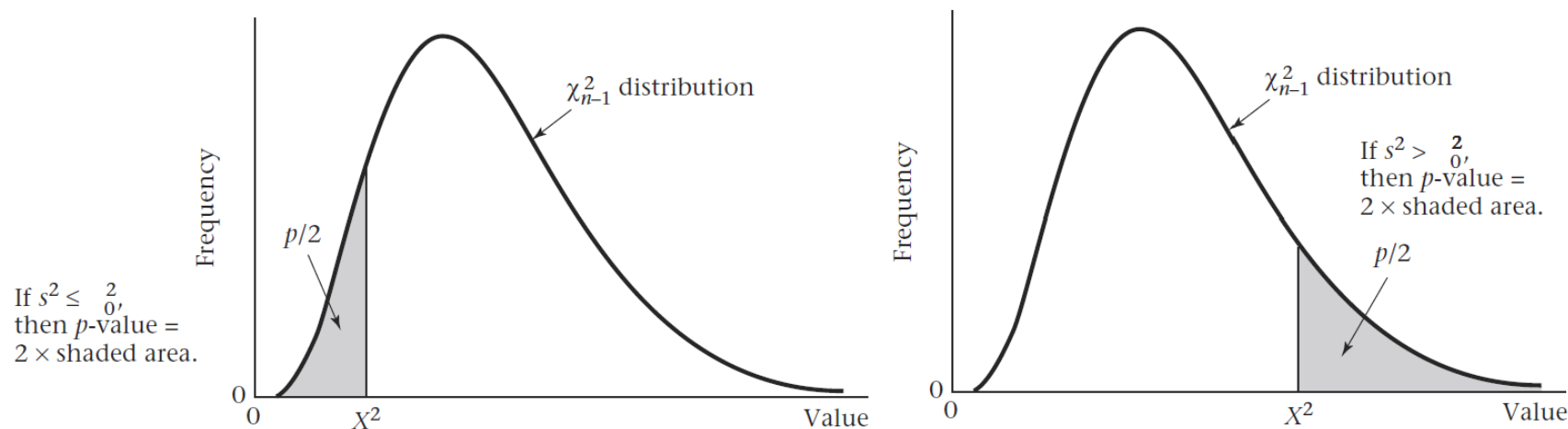
FIGURE 7.10 Acceptance and rejection regions for the one-sample χ^2 test for the variance of a normal distribution (two-sided alternative)



p -Value for a One-Sample χ^2 Test for the Variance of a Normal Distribution (Two-Sided Alternative)

- Let the test statistic $X^2 = \frac{(n-1)s^2}{\sigma_0^2}$
- If $s^2 \leq \sigma_0^2$, then p -value = $2 \times$ (area to the left of X^2 under a χ_{n-1}^2 distribution).
- If $s^2 > \sigma_0^2$, then p -value = $2 \times$ (area to the right of X^2 under a χ_{n-1}^2 distribution).

FIGURE 7.11 Illustration of the p -value for a one-sample χ^2 test for the variance of a normal distribution (two-sided alternative)



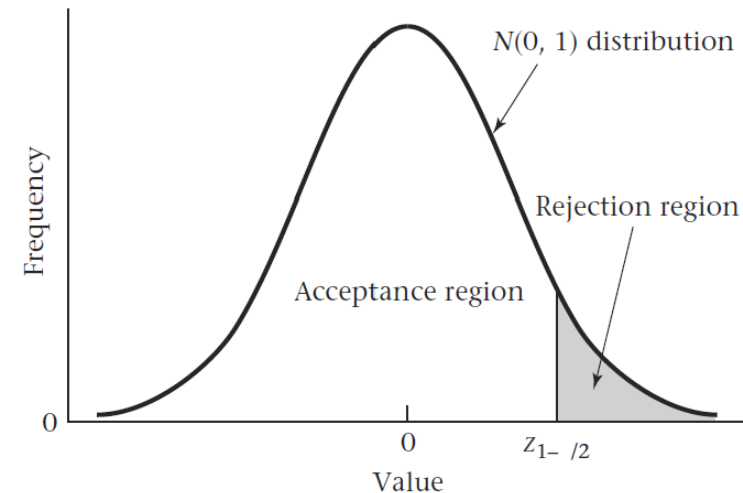
One-sample Inference for the Binomial Distribution: Normal-Theory Methods

- Let the test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

- If $z < z_{\alpha/2}$ or $z > z_{1-\alpha/2}$, then H_0 is rejected.
- If $z_{\alpha/2} \leq z \leq z_{1-\alpha/2}$, then H_0 is accepted.
- This test should only be used if $np_0 q_0 \geq 5$.

FIGURE 7.12 Acceptance and rejection regions for the one-sample binomial test—normal-theory method (two-sided alternative)



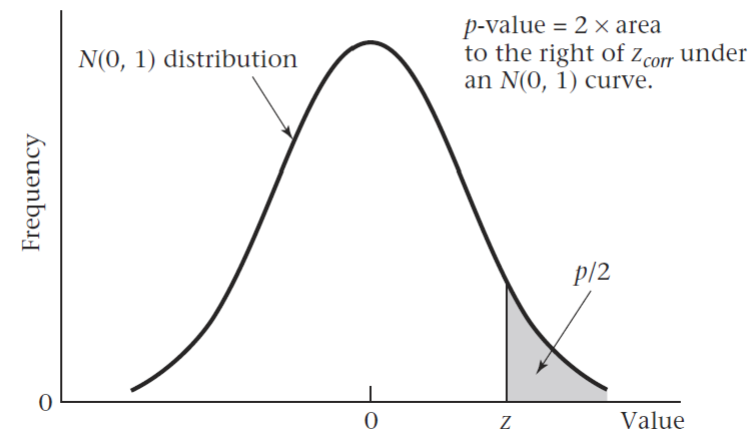
Computation of the p-value for the one-Sample Binomial Test: Normal Theory Method (Two-Sided Alternative)

- Let the test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

- If $\hat{p} \leq p_0$, then p -value = $2 \times \Phi(z)$ = twice the area to the left of z under an $N(0,1)$ curve.
- If $\hat{p} > p_0$, then p -value = $2 \times [1 - \Phi(z)]$ = twice the area to the right of z under an $N(0,1)$ curve.

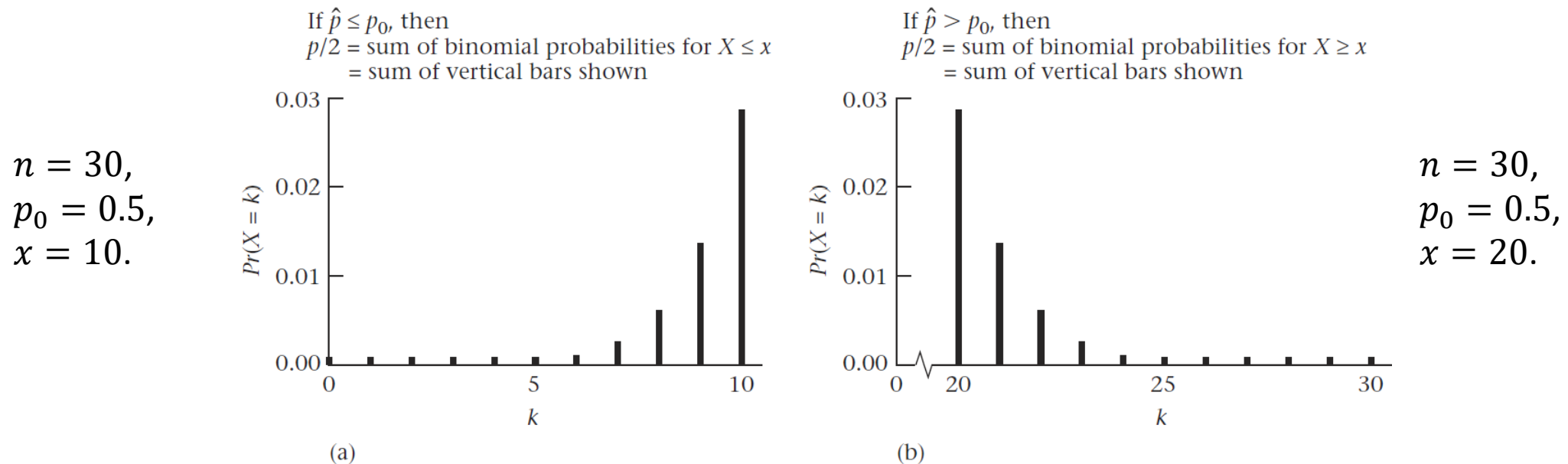
FIGURE 7.13 Illustration of the p -value for a one-sample binomial test—normal-theory method (two-sided alternative)



Computation of the p -Value for the One-Sample Binomial Test: Exact Method (Two-Sided Alternative)

- If $\hat{p} \leq p_0$, $p = 2 \times \Pr(X \leq x) = \min\left[2 \sum_{k=0}^x \binom{n}{k} p_0^k (1 - p_0)^{n-k}, 1\right]$
- If $\hat{p} > p_0$, $p = 2 \times \Pr(X \geq x) = \min\left[2 \sum_{k=x}^n \binom{n}{k} p_0^k (1 - p_0)^{n-k}, 1\right]$

FIGURE 7.14 Illustration of the p -value for a one-sample binomial test—exact method (two-sided alternative)



Power for the One-Sample Binomial Test (Two-Sided Alternative)

- $H_0: p = p_0$ vs. $H_1: p \neq p_0$, for the specific alternative $p = p_1$, the power is given by

$$\Phi \left[\sqrt{\frac{p_0 q_0}{p_1 q_1}} \left(z_{\alpha/2} + \frac{|p_0 - p_1| \sqrt{n}}{\sqrt{p_0 q_0}} \right) \right]$$

Sample-Size Estimation for the One-Sample Binomial Test (Two-Sided Alternative)

- Suppose we wish to test $H_0: p = p_0$ vs. $H_1: p \neq p_0$. The sample size needed to conduct a two-sided test with significance level α and power $1 - \beta$ vs. the specific alternative hypothesis $p = p_1$ is

$$n = \frac{p_0 q_0 \left(z_{1-\alpha/2} + z_{1-\beta} \sqrt{\frac{p_1 q_1}{p_0 q_0}} \right)^2}{(p_1 - p_0)^2}$$

One-Sample Inference for the Poisson Distribution (Small-Sample Test: Critical-Value Method)

- Let X be a Poisson random variable with expected value $= \mu$. To test the hypothesis $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ using a two-sided test with significance level α ,
- Obtain the two-sided $100\% \times (1 - \alpha)$ CI for μ based on the observed value x of X . Denote this CI (c_1, c_2) .
- If $\mu_0 < c_1$ or $\mu_0 > c_2$, then reject H_0 .
- If $c_1 \leq \mu_0 \leq c_2$, then accept H_0 .

One-Sample Inference for the Poisson Distribution (Small-Sample Test: p -Value Method)

Let μ = expected value of a Poisson distribution. To test the hypothesis $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$.

1. Compute x = observed number of deaths in the study population.
2. Under H_0 , the random variable X will follow a Poisson distribution with parameter μ_0 . Thus, the exact two-sided p -value is given by

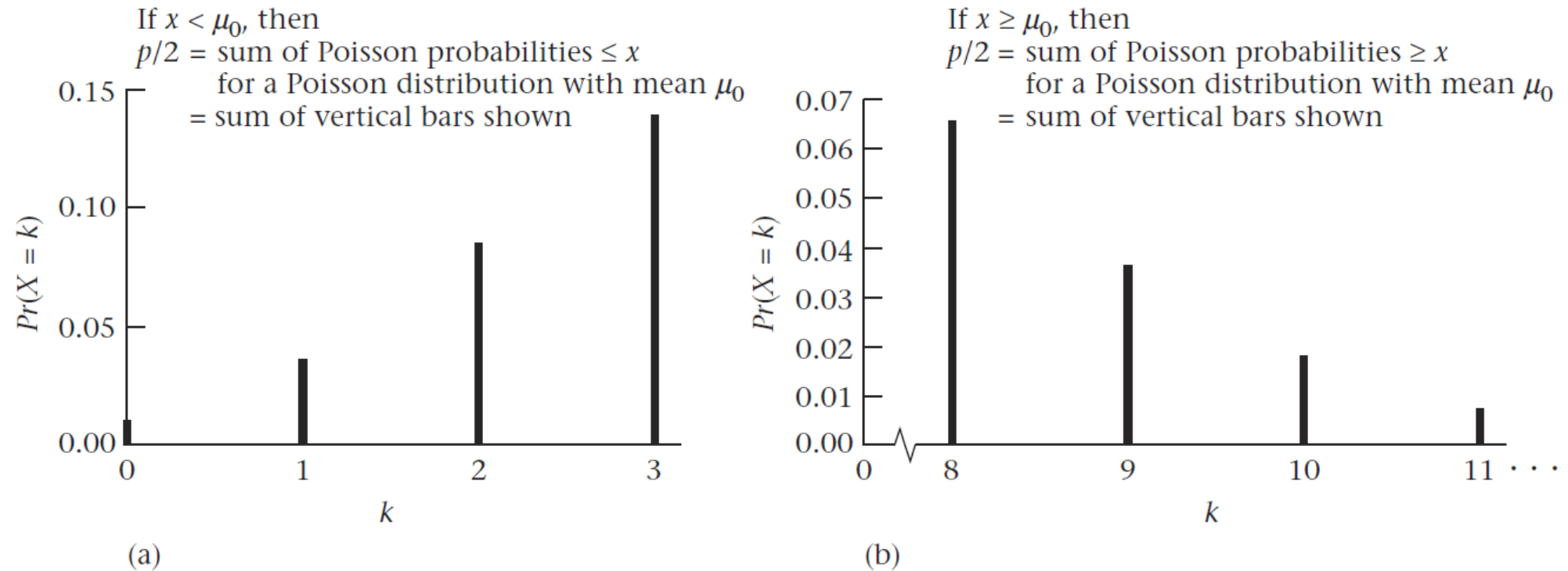
- If $x < \mu_0$:

$$p = \min \left(2 \times \sum_{k=0}^x \frac{e^{-\mu_0} \mu_0^k}{k!}, 1 \right)$$

- If $x \geq \mu_0$:

$$p = \min \left[2 \times \left(1 - \sum_{k=0}^{x-1} \frac{e^{-\mu_0} \mu_0^k}{k!} \right), 1 \right]$$

FIGURE 7.15 Computation of the exact p -value for the one-sample Poisson test



One-Sample Inference for the Poisson Distribution (Large-Sample Test)

- Let μ = expected value of a Poisson random variable. To test the hypothesis $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$,
- Compute x = observed number of events in the study population.
- Compute the test statistic

$$X^2 = \frac{(x - \mu_0)^2}{\mu_0} = \mu_0 \left(\frac{x}{\mu_0} - 1 \right)^2 \sim \chi_1^2 \text{ under } H_0$$

- For a two-sided test at level α , H_0 is rejected if $X^2 > \chi_{1,1-\alpha}^2$ and H_0 is accepted if $X^2 \leq \chi_{1,1-\alpha}^2$.
- The exact p -value is approximately given by $\Pr(\chi_1^2 > X^2)$
- This test should only be used if $\mu_0 \geq 10$.
- An approximate $100\% \times (1 - \alpha)$ CI for μ is given by $x \pm z_{1-\alpha/2} \sqrt{x}$

FIGURE 7.16 Acceptance and rejection regions for the one-sample Poisson test (large-sample test)

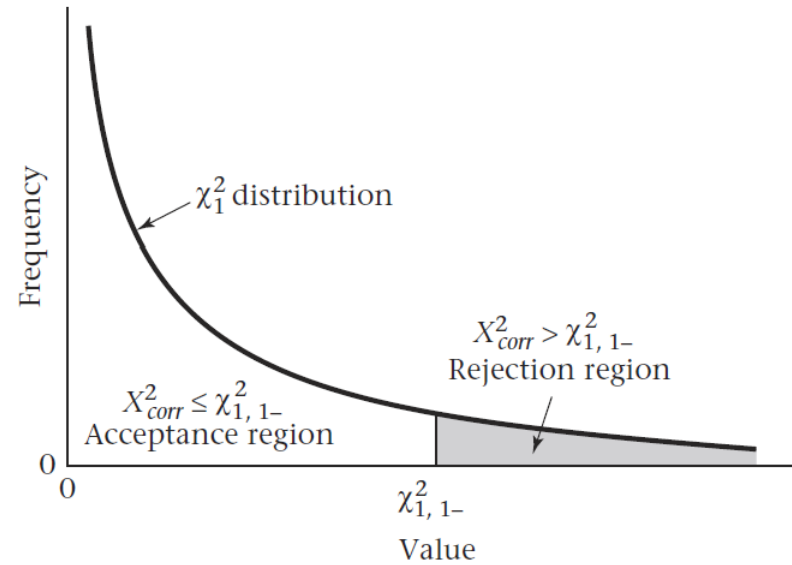
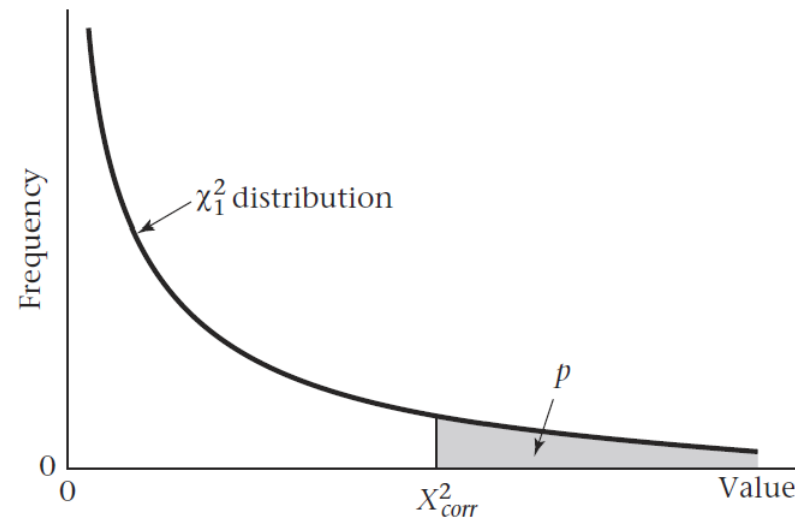


FIGURE 7.17 Computation of the p -value for the one-sample Poisson test (large-sample test)



Summary

- In this chapter, we introduced
 1. Specification of the **null (H_0)** and **alternative (H_1)** hypotheses;
 2. **type I error (α)**, **type II error (β)**, and **the power ($1-\beta$)** of a hypothesis test; the p-value of a hypothesis test and the distinction between one-sided and two-sided tests;
 3. methods for estimating appropriate **sample size** as determined by the prespecified null and alternative hypotheses and type I and type II errors.
- These concepts were applied to many one-sample hypothesis-testing cases. Each of the hypothesis tests was shown to be conducted in one of two ways
 1. Specifying critical values to determine the acceptance and rejection regions (**critical-value method**) based on a specified type I error α .
 2. Computing p-values (**p-value method**)
- **Bayesian inference methods** (a) when no prior information exists concerning a parameter and (b) when a substantial amount of prior information is available.

Key Formulas for Hypothesis Testing (One-Sample)

Key Concept	Formula
One-Sample z Test	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$
One-Sample t Test	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$
Chi-Square Variance Test	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi_{n-1}^2$
Binomial Proportion Test	$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$
Poisson Test	$\chi^2 = \frac{(x - \mu_0)^2}{\mu_0} \sim \chi_1^2$
Two-Sided p -Value (z Test)	$p = 2[1 - \Phi(z)]$
Sample Size (Two-Sided z Test)	$n \approx \frac{\sigma^2(z_{1-\alpha/2} + z_{1-\beta})^2}{(\mu_1 - \mu_0)^2}$