

2. Tests in the Normal Model

Preliminaries

The Normal Model

Suppose that $X = (X_1, X_2, \dots, X_n)$ is a [random sample](#) of size n from the [normal distribution](#) with [mean](#) $\mu \in \mathbb{R}$ and [standard deviation](#) $\sigma \in (0, \infty)$. In general, both parameters are unknown, so the parameter space is $\mathbb{R} \times (0, \infty)$.

In this section we will construct [hypothesis tests](#) for μ and σ . among the most important special cases of hypothesis testing. We will develop several tests, based on different test statistics. Some of these tests will work better than others, depending on whether one of the parameters is known. This section parallels the section on [Estimation in the Normal Model](#) in the chapter on [Set Estimation](#), and in particular, the duality between set estimation and hypothesis testing will play an important role.

The Basic Test Statistics

Ultimately, all of our test statistics can be constructed from basic test statistics that are “standardized” versions of our data variables. For $a \in \mathbb{R}$ and $b \in (0, \infty)$ let

$$Z_i(a, b) = \frac{X_i - a}{b}, \quad i \in \{1, 2, \dots, n\}$$

1. Show that $Z(a, b) = (Z_1(a, b), Z_2(a, b), \dots, Z_n(a, b))$ is a random sample of size n from the normal distribution with mean $\frac{\mu - a}{b}$ and standard deviation $\frac{\sigma}{b}$

In particular, $Z(\mu, \sigma)$ is a random sample from the standard normal distribution. Thus, it is reasonable that test statistics constructed from $Z(a, b)$ may be useful for testing hypothesis with a and b as conjectured values of μ and σ respectively.

Tests Based on a Normal Statistic

The Test Statistic

Recall that the [sample mean](#) of our data vector is

$$M = \frac{1}{n} \sum_{i=1}^n X_i$$

Our first test statistic is a linear transformation of the sample mean. For $a \in \mathbb{R}$ and $b \in (0, \infty)$, define

$$Z(a, b) = \frac{M - a}{b / \sqrt{n}}$$

2. Show that $Z(a, b)$ has the normal distribution with mean $\frac{\mu - a}{b / \sqrt{n}}$ and standard deviation $\frac{\sigma}{b}$

In particular, the random variable $Z(\mu, \sigma)$ is the ordinary *standard score* of M and has the standard normal distribution; this variable was one of the **pivot variables** used to construct [confidence sets in the normal model](#).

3. Show that the test statistic in [Exercise 2](#) can be written in terms of the basic test statistics in [Exercise 1](#) as follows:

$$Z(a, b) = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i(a, b)$$

As usual, we will let ϕ denote the standard normal probability density function and Φ the standard normal distribution function. For $p \in (0, 1)$, let $z(p)$ denote the [quantile](#) of order p for the standard normal distribution. That is, $z(p) = \Phi^{-1}(p)$. For selected values of p , $z(p)$ can be obtained from the last row of the [table of the \$t\$ distribution](#), from the [table of the standard normal distribution](#), from the [quantile applet](#), or from most statistical software packages.

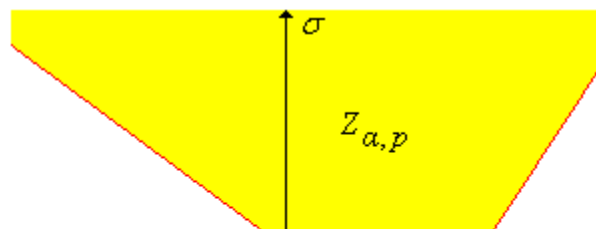
4. Show or recall the following properties. Part (d) follows from the inverse function theorem of calculus.

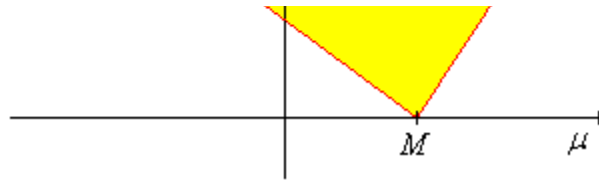
- $z(p) = -z(1 - p)$
- $z(p) \downarrow -\infty$ as $p \downarrow 0$
- $z(p) \uparrow \infty$ as $p \uparrow 1$
- $z'(p) = \frac{1}{\phi(z(p))}$

Hypothesis Tests

5. Show that for any $\alpha \in (0, 1)$ and $p \in (0, 1)$, the following test has significance level α . Reject $H_0: (\mu, \sigma) = (\mu_0, \sigma_0)$ versus $H_1: (\mu, \sigma) \neq (\mu_0, \sigma_0)$ if and only if $Z(\mu_0, \sigma_0) \leq z(\alpha - p\alpha)$ or $Z(\mu_0, \sigma_0) \geq z(1 - p\alpha)$. Equivalently, we *fail* to reject H_0 if and only if

$$M - z(1 - p\alpha) \frac{\sigma_0}{\sqrt{n}} < \mu_0 < M - z(\alpha - p\alpha) \frac{\sigma_0}{\sqrt{n}}$$





Note that the hypothesis test in [Exercise 5](#) is the dual of the confidence set constructed with the pivot variable $Z(\mu, \sigma)$ in the section on [Estimation in the Normal Model](#). That is, the set of (μ_0, σ_0) for which we fail to reject H_0 at level α is precisely the $1 - \alpha$ confidence set for (μ, σ) . This set is shown in the picture above. Note that p is the fraction of the significance level α in the right tail of the distribution of the pivot variable; $1 - p$ is the fraction of α in the left tail. The **equal-tailed** case when $p = \frac{1}{2}$ is the one that is most commonly used; this test is said to be **unbiased**.

Clearly, a test of two real parameters based on a single real-valued test statistic cannot be very good. In the remainder of this subsection, we will assume that σ is known. This is often, but not always, an artificial assumption; see [Exercise 43](#) for an example where the assumption may be reasonable. With σ known, we will write out basic test statistic as $Z(\mu_0) = Z(\mu_0, \sigma)$.

6. Show that for any $\alpha \in (0, 1)$ and $p \in (0, 1)$, the following test has significance level α :
- Reject $H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$ if and only if $Z(\mu_0) \leq z(\alpha - p\alpha)$ or $Z(\mu_0) \geq z(1 - p\alpha)$
 - Reject $H_0: \mu \geq \mu_0$ versus $H_1: \mu < \mu_0$ if and only if $Z(\mu_0) \leq z(\alpha) = -z(1 - \alpha)$
 - Reject $H_0: \mu \leq \mu_0$ versus $H_1: \mu > \mu_0$ if and only if $Z(\mu_0) \geq z(1 - \alpha)$

In part (a), the unbiased test with $p = \frac{1}{2}$ is most commonly used: Reject $H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$ if and only if $Z(\mu_0) \leq -z(1 - \frac{\alpha}{2})$ or $Z(\mu_0) \geq z(1 - \frac{\alpha}{2})$

7. For each of the tests in [Exercise 6](#), show that we fail to reject H_0 at significance level α if and only if μ_0 is in the corresponding $1 - \alpha$ confidence interval.

Power Functions

Recall that the **power function** of a test of a parameter is the probability of rejecting the null hypothesis, as a function of the true value of the parameter. Our next series of exercises will explore the power functions of the tests in [Exercise 6](#).

8. Show that the power function of the test in [Exercise 6 \(a\)](#) is given by the formula below, and satisfies the given properties
- $$Q(\mu) = \Phi\left(z(\alpha - p\alpha) - \frac{\sqrt{n}}{\sigma}(\mu - \mu_0)\right) + \Phi\left(\frac{\sqrt{n}}{\sigma}(\mu - \mu_0) - z(1 - p\alpha)\right)$$

- Q is decreasing on $(-\infty, m_0)$ and increasing on (m_0, ∞) where $m_0 = \mu_0 + (z(\alpha - p\alpha) + z(1 - p\alpha)) \frac{\sqrt{n}}{2\sigma}$.
- $Q(\mu_0) = \alpha$.
- $Q(\mu) \rightarrow 1$ as $\mu \uparrow \infty$ and $Q(\mu) \rightarrow 0$ as $\mu \downarrow -\infty$.
- If $p = \frac{1}{2}$ then Q is symmetric about μ_0 .
- As p increases, $Q(\mu)$ increases if $\mu > \mu_0$ and decreases if $\mu < \mu_0$.

9. Show that the power function of the test in [Exercise 6 \(b\)](#), is given by the formula below, and satisfies the given properties:

$$Q(\mu) = \Phi\left(z(\alpha) + \frac{\sqrt{n}}{\sigma}(\mu - \mu_0)\right)$$

- Q is increasing on \mathbb{R} .
- $Q(\mu_0) = \alpha$.
- $Q(\mu) \rightarrow 1$ as $\mu \uparrow \infty$ and $Q(\mu) \rightarrow 0$ as $\mu \downarrow -\infty$.

10. Show that the power function of the test in [Exercise 6 \(c\)](#), is given by the formula below, and satisfies the given properties:

$$Q(\mu) = \Phi\left(z(\alpha) - \frac{\sqrt{n}}{\sigma}(\mu - \mu_0)\right)$$

- Q is decreasing on \mathbb{R} .
- $Q(\mu_0) = \alpha$.
- $Q(\mu) \rightarrow 0$ as $\mu \uparrow \infty$ and $Q(\mu) \rightarrow 1$ as $\mu \downarrow -\infty$.

11. Show that for any of the three tests in [Exercise 6](#), increasing the sample size n or decreasing the standard deviation σ results in a uniformly more powerful test.

12. In the **mean test experiment**, select the normal test statistic and select the normal sampling distribution with standard deviation $\sigma = 2$, significance level $\alpha = 0.1$, sample size $n = 20$, and $\mu_0 = 0$. Run the experiment 1000 times, updating every 10 runs for several values of the true distribution mean μ . For each value of μ , note the relative frequency of the event that the null hypothesis is rejected. Sketch the empirical power function.

13. In the **mean estimate experiment**, select the normal pivot variable and select the normal distribution with $\mu = 0$ and standard deviation $\sigma = 2$, confidence level $1 - \alpha = 0.90$, and sample size $n = 10$. For each of the three types of confidence intervals, run the experiment 20 times, updating after each run. State the corresponding hypotheses and significance level, and for each run, give the set of μ_0 for which the null hypothesis would be rejected.

Design of the Experiment

In many cases, the first step is to *design* the experiment so that the significance level is α and so that the test has a given power β for a given alternative μ_1

14. For either of the one-sided tests in [Exercise 6](#), show that the sample size n needed for a test with significance level α and power β for the alternative μ_1 is as follows. *Hint*: Set the power function equal to β and solve for n

$$n = \left(\frac{\sigma(z(\beta) - z(\alpha))}{\mu_1 - \mu_0} \right)^2$$

15. For the unbiased, two-sided test, show that the sample size n needed for a test with significance level α and power β for the alternative μ_1 is given approximately by the following formula. *Hint*: In the power function for the two-sided test given in [Exercise 8](#), neglect the first term if $\mu_1 < \mu_0$ and neglect the second term if $\mu_1 > \mu_0$.

$$n = \left(\frac{\sigma(z(\beta) - z(\alpha/2))}{\mu_1 - \mu_0} \right)^2$$

Tests Based on Chi-Square Statistics

Test Statistics

In this section we will use test statistics related to the [sample variance](#). First recall that the usual version of the sample variance is

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - M)^2$$

Moreover, recall that one of the most important [special properties of normal samples](#) is that the sample mean M and the sample variance S^2 are independent. Now, for $a \in \mathbb{R}$ and $b \in (0, \infty)$, define

$$W^2(a) = \frac{1}{n} \sum_{i=1}^n (X_i - a)^2, \quad U(a, b) = \frac{n}{b^2} W^2(a), \quad V(b) = \frac{n-1}{b^2} S^2$$

Recall that we have used $W^2(\mu)$ as a special version of the sample variance, in the unlikely event that the mean μ is known.

16. Show that $U(a, b)$ is the sum of the squares of the basic test statistics in [Exercise 1](#):

$$U(a, b) = \sum_{i=1}^n Z_i^2(a, b)$$

17. Show that $V(b) = U(a, b) - Z^2(a, b)$ where $Z(a, b)$ is the test statistic in [Exercise 2](#). Thus, in particular, it follows that $V(b)$ can also be written in terms of the basic test statistics in [Exercise 1](#).

18. Show that

- $U(\mu, \sigma)$ has the chi-square distribution with n degrees of freedom.
- $V(\sigma)$ has the chi-square distribution with $n - 1$ degrees of freedom.

The variables in Exercise 18 are **pivot variables** that were used to construct [confidence sets in the normal model](#).

Let g_k denote the probability density function and G_k the distribution function for the [chi-square distribution](#) with k degrees of freedom. In addition, for $p \in (0, 1)$, let $\chi_k^2(p)$ denote the [quantile](#) of order p for the distribution, so that by definition, $\chi_k^2(p) = G_k^{-1}(p)$. For selected values of k and p , $\chi_k^2(p)$ can be obtained from the [table of the chi-square distribution](#), from the [quantile applet](#), or from most statistical software packages.

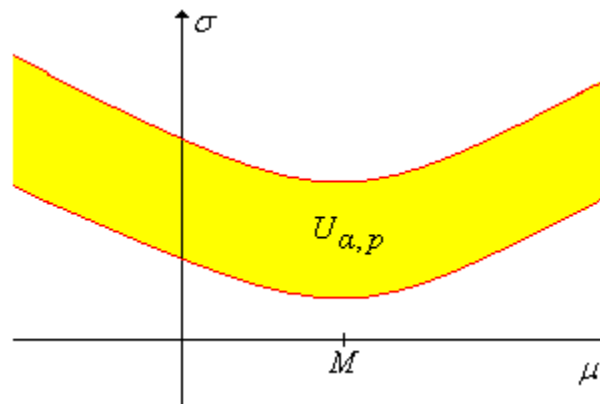
19. Show or recall the following properties. Part (c) follows from the inverse function theorem of calculus:

- $\chi_k^2(p) \rightarrow 0$ as $p \downarrow 0$
- $\chi_k^2(p) \rightarrow \infty$ as $p \uparrow 1$
- $\frac{d\chi_k^2(p)}{dp} = \frac{1}{g_k(\chi_k^2(p))}$

Hypothesis Tests

20. Show that for any $\alpha \in (0, 1)$ and $p \in (0, 1)$, the following test has significance level α . Reject $H_0: (\mu, \sigma) = (\mu_0, \sigma_0)$ versus $H_1: (\mu, \sigma) \neq (\mu_0, \sigma_0)$ if and only if $U(\mu_0, \sigma_0) \leq \chi_n^2(\alpha - p\alpha)$ or $U(\mu_0, \sigma_0) \geq \chi_n^2(1 - p\alpha)$. Equivalently, show that we *fail* to reject H_0 is and only if

$$\frac{n W^2(\mu_0)}{\chi_n^2(1 - p\alpha)} < \sigma_0^2 < \frac{n W^2(\mu_0)}{\chi_n^2(\alpha - p\alpha)}$$

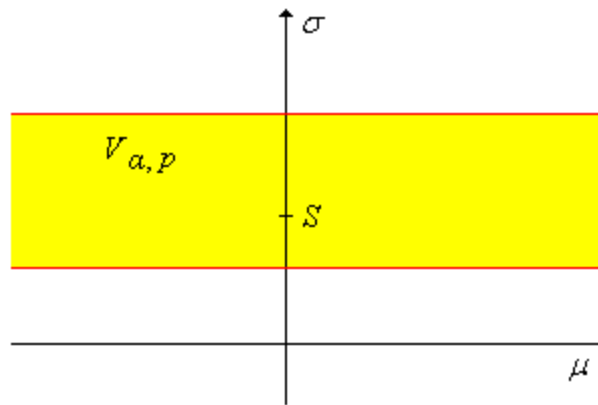


Note that the hypothesis test in [Exercise 20](#) is the dual of the confidence sets constructed with the pivot variable $U(\mu, \sigma)$ in the section on [Estimation in the Normal Model](#). That is, the set of (μ_0, σ_0) for which we

fail to reject H_0 at level α is precisely the $1 - \alpha$ confidence set for (μ, σ) . This set are shown in the pictures above.

21. Show that for any $\alpha \in (0, 1)$ and $p \in (0, 1)$, the following test has significance level α . Reject $H_0: (\mu, \sigma) = (\mu_0, \sigma_0)$ versus $H_1: (\mu, \sigma) \neq (\mu_0, \sigma_0)$ if and only if $V(\sigma_0) \leq \chi_{n-1}^2(\alpha - p\alpha)$ or $V(\sigma_0) \geq \chi_{n-1}^2(1 - p\alpha)$. Equivalently, show that we *fail* to reject H_0 if and only if

$$\frac{(n-1)S^2}{\chi_{n-1}^2(1-p\alpha)} < \sigma_0^2 < \frac{(n-1)S^2}{\chi_{n-1}^2(\alpha-p\alpha)}$$



Note that the hypothesis tests in [Exercise 21](#) is the dual of the confidence set constructed with the pivot variable $V(\sigma)$ in the section on [Estimation in the Normal Model](#). That is, the set of (μ_0, σ_0) for which we fail to reject H_0 at level α is precisely the $1 - \alpha$ confidence set for (μ, σ) . This set is shown in the picture above.

In both tests, note that p is the fraction of the significance level α in the right tail of the distribution of the pivot variable; $1 - p$ is the fraction of α in the left tail. The **equal-tailed** case when $p = \frac{1}{2}$ is the one that is most commonly used; this test is said to be **unbiased**.

Clearly, a test of two real parameters based on a single real-valued test statistic cannot be very good. However, since the test statistic $V(\sigma_0)$ gives no information about μ , it is natural to use this statistic in tests involving σ only. If μ were known, we could also use the test statistic $U(\sigma_0) = U(\mu, \sigma_0)$. All of the results given in the remainder of this subsection would apply, but with n replacing $n - 1$. However, the assumption that μ is known is very artificial.

22. Show that for any $\alpha \in (0, 1)$ and $p \in (0, 1)$, the following test has significance level α :

- Reject $H_0: \sigma = \sigma_0$ versus $H_1: \sigma \neq \sigma_0$ if and only if $V(\sigma_0) \leq \chi_{n-1}^2(\alpha - p\alpha)$ or $V(\sigma_0) \geq \chi_{n-1}^2(1 - p\alpha)$
- Reject $H_0: \sigma \geq \sigma_0$ versus $H_1: \sigma < \sigma_0$ if and only if $V(\sigma_0) \leq \chi_{n-1}^2(\alpha)$
- Reject $H_0: \sigma \leq \sigma_0$ versus $H_1: \sigma > \sigma_0$ if and only if $V(\sigma_0) \geq \chi_{n-1}^2(1 - \alpha)$

In part (a), the unbiased test with $p = \frac{1}{2}$ is most commonly used: Reject $H_0: \sigma = \sigma_0$ versus $H_1: \sigma \neq \sigma_0$ if

and only if $V(\sigma_0) \leq \chi_{n-1}^2\left(\frac{\alpha}{2}\right)$ or $V(\sigma_0) \geq \chi_{n-1}^2\left(1 - \frac{\alpha}{2}\right)$

23. For each of the tests in [Exercise 22](#), show that we fail to reject H_0 at significance level α if and only if σ_0 is in the corresponding $1 - \alpha$ confidence interval.

Power Functions

Recall that the **power function** of a test of a parameter is the probability of rejecting the null hypothesis, as a function of the true value of the parameter. Our next series of exercises will explore the power functions of the tests in [Exercise 22](#).

24. Show that the power function of the test in [Exercise 22 \(a\)](#) is given by the following formula, and then verify the given properties:

$$Q(\sigma) = 1 - G_{n-1}\left(\frac{\sigma_0^2}{\sigma^2} \chi_{n-1}^2(1 - p\alpha)\right) + G_{n-1}\left(\frac{\sigma_0^2}{\sigma^2} \chi_{n-1}^2(\alpha - p\alpha)\right)$$

- Q is decreasing on $(-\infty, \sigma_0)$ and increasing on (σ_0, ∞) .
 - $Q(\sigma_0) = \alpha$.
 - $Q(\sigma) \rightarrow 1$ as $\sigma \uparrow \infty$ and $Q(\sigma) \rightarrow 1$ as $\sigma \downarrow 0$
25. Show that the power function of the test in [Exercise 22 \(b\)](#) is given by the following formula, and then verify the given properties:

$$Q(\sigma) = 1 - G_{n-1}\left(\frac{\sigma_0^2}{\sigma^2} \chi_{n-1}^2(1 - \alpha)\right)$$

- Q is increasing on $(0, \infty)$
- $Q(\sigma_0) = \alpha$.
- $Q(\sigma) \rightarrow 1$ as $\sigma \uparrow \infty$ and $Q(\sigma) \rightarrow 0$ as $\sigma \downarrow 0$

26. Show that the power function for the test in [Exercise 22 \(c\)](#) is given by the following formula, and then verify the given properties:

$$Q(\sigma) = G_{n-1}\left(\frac{\sigma_0^2}{\sigma^2} \chi_{n-1}^2(\alpha)\right)$$

- Q is decreasing on $(0, \infty)$
- $Q(\sigma_0) = \alpha$.
- $Q(\sigma) \rightarrow 0$ as $\sigma \uparrow \infty$ and $Q(\sigma) \rightarrow 1$ as $\sigma \downarrow 0$

27. Show that for any of the three tests in [Exercise 22](#), increasing the sample size n results in a uniformly more powerful test.

28. In the **variance test experiment**, select the normal distribution with mean 0, and select significance level 0.1, sample size 10, and test standard deviation 1.0. For various values of the true standard deviation, run the simulation 1000 times, updating every 10 runs. Record the relative frequency of rejecting the null hypothesis and plot the empirical power curve.

- Two-sided test
- Left-tailed test
- Right-tailed test

29. In the **variance estimate experiment**, select the normal distribution with mean 0 and standard deviation 2, and select confidence level 0.90 and sample size 10. Run the experiment 20 times, updating after each run. State the corresponding hypotheses and significance level, and for each run, give the set of test standard deviations for which the null hypothesis would be rejected.

- Two-sided confidence interval
- Confidence lower bound
- Confidence upper bound

Tests Based on a Student t Statistic

The Test Statistic

Our next test statistic will lead to good tests of μ without requiring the assumption that σ is known. For $a \in \mathbb{R}$, define

$$T(a) = \frac{M - a}{S / \sqrt{n}}$$

30. Show that for any $a \in \mathbb{R}$ and any $b \in (0, \infty)$,

$$T(a) = \frac{Z(a, b)}{\sqrt{\frac{V(b)}{n-1}}}$$

In particular, it follows that $T(a)$ can be written in terms of the basic test statistics in [Exercise 1](#).

31. Show or recall that $T(\mu)$ has the **student t distribution** with $n - 1$ degrees of freedom.

This variable was one of the pivot variables used to construct [estimates in the normal model](#). When $\mu \neq a$, the distribution of $T(a)$ is known as a **non-central t distribution**.

As usual, for $k > 0$, we will let ϕ_k denote the probability density function and Φ_k the distribution function for the t distribution with k degrees of freedom. Additionally, for $p \in (0, 1)$, let $t_k(p)$ denote the **quantile** of order p for this distribution, so that $t_k(p) = \Phi_k^{-1}(p)$. For selected values of k and p , values of $t_k(p)$ can be

obtained from the [table of the \$t\$ distribution](#) or from the [quantile applet](#).

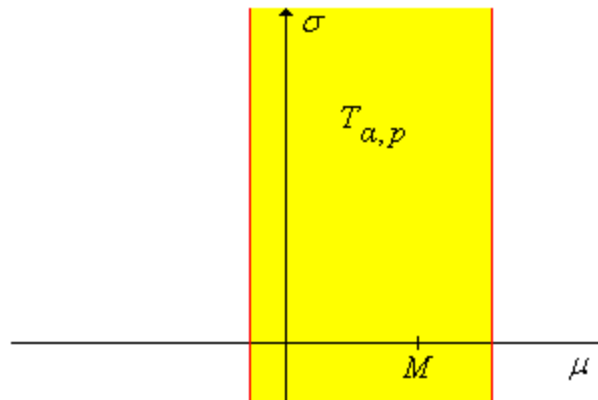
32. Show or recall the following properties. Part (d) follows from the inverse function theorem of calculus.

- $t_k(1 - p) = -t_k(p)$
- $t_k(p) \rightarrow -\infty$ as $p \downarrow 0$
- $t_k(p) \rightarrow \infty$ as $p \uparrow 1$
- $t_k'(p) = \frac{1}{\phi_k(t_k(p))}$

Hypothesis Tests

33. Show that for any $\alpha \in (0, 1)$ and $p \in (0, 1)$, the following test has significance level α . Reject $H_0: (\mu, \sigma) = (\mu_0, \sigma_0)$ versus $H_1: (\mu, \sigma) \neq (\mu_0, \sigma_0)$ if and only if $T(\mu_0) \leq t_{n-1}(\alpha - p\alpha)$ or $T(\mu_0) \geq t_{n-1}(1 - p\alpha)$. Equivalently, we *fail* to reject H_0 if and only if

$$M - t_{n-1}(1 - p\alpha) \frac{S}{\sqrt{n}} < \mu_0 < M - t_{n-1}(\alpha - p\alpha) \frac{S}{\sqrt{n}}$$



Note that the hypothesis test in [Exercise 33](#) is the dual of the confidence set constructed with the pivot variable $T(\mu)$ in the section on [Estimation in the Normal Model](#). That is, the set of (μ_0, σ_0) for which we fail to reject H_0 at level α is precisely the $1 - \alpha$ confidence set for (μ, σ) . This set is shown in the picture above. Note that p is the fraction of the significance level α in the right tail of the distribution of the pivot variable; $1 - p$ is the fraction of α in the left tail. The **equal-tailed** case when $p = \frac{1}{2}$ is the one that is most commonly used; this test is said to be **unbiased**.

Clearly since our test statistic gives no information about σ , there is no point in putting this parameter in the hypotheses.

34. Show that for any $\alpha \in (0, 1)$ and $p \in (0, 1)$, the following test has significance level α :

- Reject $H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$ if and only if $T(\mu_0) \leq t_{n-1}(\alpha - p\alpha)$ or $T(\mu_0) \geq t_{n-1}(1 - p\alpha)$
- Reject $H_0: \mu \geq \mu_0$ versus $H_1: \mu < \mu_0$ if and only if $T(\mu_0) \leq t_{n-1}(\alpha) = -t_{n-1}(1 - \alpha)$

c. Reject $H_0: \mu \leq \mu_0$ versus $H_1: \mu > \mu_0$ if and only if $T(\mu_0) \geq t_{n-1}(1 - \alpha)$

In part (a), the unbiased test with $p = \frac{1}{2}$ is most commonly used:

Reject $H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$ if and only if $T(\mu_0) \leq -t_{n-1}(1 - \frac{\alpha}{2})$ or $T(\mu_0) \geq t_{n-1}(1 - \frac{\alpha}{2})$

35. For each of the tests in [Exercise 34](#), show that we fail to reject H_0 at significance level α if and only if μ_0 is in the corresponding $1 - \alpha$ confidence interval.

The P -value of these test can be computed in terms of the distribution function Φ_{n-1} of the t -distribution with $n - 1$ degrees of freedom.

36. Show that the P -values of the tests in [Exercise 34](#) are respectively

- $2(1 - \Phi_{n-1}(|T_0|))$
- $1 - \Phi_{n-1}(T_0)$
- $\Phi_{n-1}(Z_0)$

37. In the [mean test experiment](#), select the student test statistic and select the normal sampling distribution with standard deviation $\sigma = 2$, significance level $\alpha = 0.1$, sample size $n = 20$, and $\mu_0 = 0$. Run the experiment 1000 times, updating every 10 runs for several values of the true distribution mean μ . For each value of μ , note the relative frequency of the event that the null hypothesis is rejected. Sketch the empirical power function.

38. In the [mean estimate experiment](#), select the student pivot variable and select the normal sampling distribution with mean 0 and standard deviation 2. Select confidence level 0.90 and sample size 10. For each of the three types of intervals, run the experiment 20 times, updating after each run. State the corresponding hypotheses and significance level, and for each run, give the set of μ_0 for which the null hypothesis would be rejected.

The power function for the tests in [Exercise 34](#) can be computed explicitly in terms of the non-central t distribution function. Qualitatively, the graphs of the power functions are similar to the case when σ is known, given in [Exercise 8](#), [Exercise 9](#), and [Exercise 10](#).

If an upper bound σ_0 on the standard deviation σ is known, then conservative estimates on the sample size needed for a given confidence level and a given margin of error can be obtained using the methods of [Exercise 14](#) and [Exercise 15](#).

Exercises

Robustness

The primary assumption that we made is that the underlying sampling distribution is normal. Of course, in real statistical problems, we are unlikely to know much about the sampling distribution, let alone whether or not it is normal. Suppose in fact that the underlying distribution is not normal. When the sample size n is relatively large, the distribution of the sample mean will still be approximately normal by the [central limit theorem](#), and thus our tests of the mean μ should still be approximately valid. On the other hand, tests of the variance σ are less robust to deviations from the assumption of normality. The following exercises explore these ideas.

❑ 39. In the [mean test experiment](#), select the [gamma distribution](#) with shape parameter 1 and scale parameter 1. For the three different tests and for various significance levels, sample sizes, and values of μ_0 , run the experiment 1000 times with an update frequency of 10. For each configuration, note the relative frequency of rejecting H_0 . When H_0 is true, compare the relative frequency with the significance level.

❑ 40. In the [mean test experiment](#), select the [uniform distribution](#) on $[0, 4]$. For the three different tests and for various significance levels, sample sizes, and values of μ_0 , run the experiment 1000 times with an update frequency of 10. For each configuration, note the relative frequency of rejecting H_0 . When H_0 is true, compare the relative frequency with the significance level.

How large n needs to be for the testing procedure to work well depends, of course, on the underlying distribution; the more this distribution deviates from normality, the larger n must be. Fortunately, convergence to normality in the central limit theorem is rapid and hence, as you observed in the exercises, we can get away with relatively small sample sizes (30 or more) in most cases.

❑ 41. In the [variance test experiment](#), select the [gamma distribution](#) with shape parameter 1 and scale parameter 1. For the three different tests and for various significance levels, sample sizes, and values of σ_0 , run the experiment 1000 times with an update frequency of 10. For each configuration, note the relative frequency of rejecting H_0 . When H_0 is true, compare the relative frequency with the significance level.

❑ 42. In the [variance test experiment](#), select the [uniform distribution](#) on $[0, 4]$. For the three different tests and for various significance levels, sample sizes, and values of μ_0 , run the experiment 1000 times with an update frequency of 10. For each configuration, note the relative frequency of rejecting H_0 . When H_0 is true, compare the relative frequency with the significance level.

Computational Exercises

❑ 43. The length of a certain machined part is supposed to be 10 centimeters. In fact, due to imperfections in the manufacturing process, the actual length is a random variable. The standard deviation is due to inherent factors in the process, which remain fairly stable over time. From historical data, the standard deviation is known with a high degree of accuracy to be 0.3. The mean, on the other hand, may be set by adjusting various parameters in the process and hence may change to an unknown value fairly frequently. We are interested in testing $H_0: \mu = 10$ versus $H_1: \mu \neq 10$.

- a. Suppose that a sample of 100 parts has mean 10.1. Perform the test at the 0.1 level of significance.

- b. Compute the P -value for the data in (a).
- c. Compute the power of the test in (a) at $\mu = 10.05$.
- d. Compute the approximate sample size needed for significance level 0.1 and power 0.8 when $\mu = 10.05$.



44. A bag of potato chips of a certain brand has an advertised weight of 250 grams. Actually, the weight (in grams) is a random variable. Suppose that a sample of 75 bags has mean 248 and standard deviation 5. At the 0.05 significance level, perform the following tests:

- a. $H_0: \mu \geq 250$ versus $H_1: \mu < 250$
- b. $H_0: \sigma \geq 7$ versus $H_1: \sigma < 7$



45. At a telemarketing firm, the length of a telephone solicitation (in seconds) is a random variable. A sample of 50 calls has mean 310 and standard deviation 25. At the 0.1 level of significance, can we conclude that

- a. $\mu > 300$?
- b. $\sigma > 20$?



46. At a certain farm the weight of a peach (in ounces) at harvest time is a random variable. A sample of 100 peaches has mean 8.2 and standard deviation 1.0. At the 0.01 level of significance, can we conclude that

- a. $\mu > 8$?
- b. $\sigma < 1.5$?



47. The hourly wage for a certain type of construction work is a random variable with standard deviation 1.25. For sample of 25 workers, the mean wage was \$6.75. At the 0.01 level of significance, can we conclude that $\mu < 7.00$?



48. Using **Michelson's data**, test to see if the velocity of light is greater than 730 (+299000) km/sec, at the 0.005 significance level.



49. Using **Cavendish's data**, test to see if the density of the earth is less than 5.5 times the density of water, at the 0.05 significance level.



50. Using **Short's data**, test to see if the parallax of the sun differs from 9 seconds of a degree, at the 0.1 significance level.



51. Using **Fisher's iris data**, perform the following tests, at the 0.1 level:

- a. The mean petal length of Setosa irises differs from 15 mm.
- b. The mean petal length of Verginica irises is greater than 52 mm.
- c. The mean petal length of Versicolor irises is less than 44 mm.



[Virtual Laboratories](#) > [9. Hypothesis Testing](#) > [1](#) **[2](#)** [3](#) [4](#) [5](#) [6](#) [7](#)

[Contents](#) | [Applets](#) | [Data Sets](#) | [Biographies](#) | [External Resources](#) | [Key words](#) | [Feedback](#) | ©