

9.2: Tests in the Normal Model

Basic Theory

The Normal Model

The normal distribution is perhaps the most important distribution in the study of mathematical statistics, in part because of the central limit theorem. As a consequence of this theorem, a measured quantity that is subject to numerous small, random errors will have, at least approximately, a normal distribution. Such variables are ubiquitous in statistical experiments, in subjects varying from the physical and biological sciences to the social sciences.

So in this section, we assume that $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a random sample from the normal distribution with mean μ and standard deviation σ . Our goal in this section is to construct hypothesis tests for μ and σ ; these are among of the most important special cases of hypothesis testing. This section parallels the section on Estimation in the Normal Model in the chapter on Set Estimation, and in particular, the duality between interval estimation and hypothesis testing will play an important role. But first we need to review some basic facts that will be critical for our analysis.

Recall that the sample mean M and sample variance S^2 are

$$M = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - M)^2 \quad (9.2.1)$$

From our study of point estimation, recall that M is an unbiased and consistent estimator of μ while S^2 is an unbiased and consistent estimator of σ^2 . From these basic statistics we can construct the test statistics that will be used to construct our hypothesis tests. The following results were established in the section on Special Properties of the Normal Distribution.

Define

$$Z = \frac{M - \mu}{\sigma/\sqrt{n}}, \quad T = \frac{M - \mu}{S/\sqrt{n}}, \quad V = \frac{n-1}{\sigma^2} S^2 \quad (9.2.2)$$

1. Z has the standard normal distribution.
2. T has the student t distribution with $n-1$ degrees of freedom.
3. V has the chi-square distribution with $n-1$ degrees of freedom.
4. Z and V are independent.

It follows that each of these random variables is a pivot variable for (μ, σ) since the distributions do not depend on the parameters, but the variables themselves functionally depend on one or both parameters. The pivot variables will lead to natural test statistics that can then be used to perform the hypothesis tests of the parameters. To construct our tests, we will need quantiles of these standard distributions. The quantiles can be computed using the special distribution calculator or from most mathematical and statistical software packages. Here is the notation we will use:

Let $p \in (0, 1)$ and $k \in \mathbb{N}_+$.

1. $z(p)$ denotes the quantile of order p for the standard normal distribution.
2. $t_k(p)$ denotes the quantile of order p for the student t distribution with k degrees of freedom.
3. $\chi_k^2(p)$ denotes the quantile of order p for the chi-square distribution with k degrees of freedom

Since the standard normal and student t distributions are symmetric about 0, it follows that $z(1-p) = -z(p)$ and $t_k(1-p) = -t_k(p)$ for $p \in (0, 1)$ and $k \in \mathbb{N}_+$. On the other hand, the chi-square distribution is not symmetric.

Tests for the Mean with Known Standard Deviation

For our first discussion, we assume that the distribution mean μ is unknown but the standard deviation σ is known. This is not always an artificial assumption. There are often situations where σ is stable over time, and hence is at least approximately known, while μ changes because of different “treatments”. Examples are given in the computational exercises below.

For a conjectured $\mu_0 \in \mathbb{R}$, define the test statistic

$$Z = \frac{M - \mu_0}{\sigma / \sqrt{n}} \quad (9.2.3)$$

1. If $\mu = \mu_0$ then Z has the standard normal distribution.
2. If $\mu \neq \mu_0$ then Z has the normal distribution with mean $\frac{\mu - \mu_0}{\sigma / \sqrt{n}}$ and variance 1.

So in case (b), $\frac{\mu - \mu_0}{\sigma / \sqrt{n}}$ can be viewed as a *non-centrality parameter*. The graph of the probability density function of Z is like that of the standard normal probability density function, but shifted to the right or left by the non-centrality parameter, depending on whether $\mu > \mu_0$ or $\mu < \mu_0$.

For $\alpha \in (0, 1)$, each of the following tests has significance level α :

1. Reject $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ if and only if $Z < -z(1 - \alpha/2)$ or $Z > z(1 - \alpha/2)$ if and only if $M < \mu_0 - z(1 - \alpha/2) \frac{\sigma}{\sqrt{n}}$ or $M > \mu_0 + z(1 - \alpha/2) \frac{\sigma}{\sqrt{n}}$.
2. Reject $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$ if and only if $Z > z(1 - \alpha)$ if and only if $M > \mu_0 + z(1 - \alpha) \frac{\sigma}{\sqrt{n}}$.
3. Reject $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0$ if and only if $Z < -z(1 - \alpha)$ if and only if $M < \mu_0 - z(1 - \alpha) \frac{\sigma}{\sqrt{n}}$.

Proof

In part (a), H_0 is a simple hypothesis, and under H_0 , Z has the standard normal distribution. So α is probability of falsely rejecting H_0 by definition of the quantiles. In parts (b) and (c), Z has a non-central normal distribution under H_0 as discussed [above](#). So if H_0 is true, the the maximum type 1 error probability α occurs when $\mu = \mu_0$. The decision rules in terms of M are equivalent to the corresponding ones in terms of Z by simple algebra.

Part (a) is the standard two-sided test, while (b) is the right-tailed test and (c) is the left-tailed test. Note that in each case, the hypothesis test is the dual of the corresponding interval estimate constructed in the section on Estimation in the Normal Model.

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

1. $M - z(1 - \alpha/2) \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq M + z(1 - \alpha/2) \frac{\sigma}{\sqrt{n}}$
2. $\mu_0 \leq M + z(1 - \alpha) \frac{\sigma}{\sqrt{n}}$
3. $\mu_0 \geq M - z(1 - \alpha) \frac{\sigma}{\sqrt{n}}$

Proof

This follows from the [previous result](#). In each case, we start with the inequality that corresponds to not rejecting H_0 and solve for μ_0 .

The two-sided test in (a) corresponds to $\alpha/2$ in each tail of the distribution of the test statistic Z , under H_0 . This set is said to be *unbiased*. But of course we can construct other biased tests by partitioning the confidence level α between the left and right tails in a non-symmetric way.

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

1. $p = \frac{1}{2}$ gives the symmetric, unbiased test.
2. $p \downarrow 0$ gives the left-tailed test.
3. $p \uparrow 1$ gives the right-tailed test.

Proof

As before H_0 is a simple hypothesis, and if H_0 is true, Z has the standard normal distribution. So the probability of falsely rejecting H_0 is α by definition of the quantiles. Parts (a)–(c) follow from properties of the standard normal quantile function.

The P -value of these test can be computed in terms of the standard normal distribution function Φ .

The P -values of the standard tests [above](#) are respectively

1. $2[1 - \Phi(|Z|)]$
2. $1 - \Phi(Z)$
3. $\Phi(Z)$

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 results will explore the power functions of the tests [above](#).

The power function of the general two-sided test [above](#) is given by

$$Q(\mu) = \Phi\left(z(\alpha - p\alpha) - \frac{\sqrt{n}}{\sigma}(\mu - \mu_0)\right) + \Phi\left(\frac{\sqrt{n}}{\sigma}(\mu - \mu_0) - z(1 - p\alpha)\right), \quad \mu \in \mathbb{R} \quad (9.2.4)$$

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

So by varying p , we can make the test more powerful for some values of μ , but only at the expense of making the test less powerful for other values of μ .

The power function of the left-tailed test [above](#) is given by

$$Q(\mu) = \Phi\left(z(\alpha) + \frac{\sqrt{n}}{\sigma}(\mu - \mu_0)\right), \quad \mu \in \mathbb{R} \quad (9.2.5)$$

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

The power function of the right-tailed test [above](#), is given by

$$Q(\mu) = \Phi\left(z(\alpha) - \frac{\sqrt{n}}{\sigma}(\mu - \mu_0)\right), \quad \mu \in \mathbb{R} \quad (9.2.6)$$

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

For any of the three tests in [above](#), increasing the sample size n or decreasing the standard deviation σ results in a uniformly more powerful test.

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 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.

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. 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.

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 .

For either of the one-sided tests in [above](#), the sample size n needed for a test with significance level α and power β for the alternative μ_1 is

$$n = \left(\frac{\sigma [z(\beta) - z(\alpha)]}{\mu_1 - \mu_0} \right)^2 \quad (9.2.7)$$

Proof

This follows from setting the power function equal to β and solving for n

For the unbiased, two-sided test, the sample size n needed for a test with significance level α and power β for the alternative μ_1 is approximately

$$n = \left(\frac{\sigma [z(\beta) - z(\alpha/2)]}{\mu_1 - \mu_0} \right)^2 \quad (9.2.8)$$

Proof

In the power function for the two-sided test given [above](#), we can neglect the first term if $\mu_1 < \mu_0$ and neglect the second term if $\mu_1 > \mu_0$.

Tests of the Mean with Unknown Standard Deviation

For our next discussion, we construct tests of μ without requiring the assumption that σ is known. And in applications of course, σ is usually unknown.

For a conjectured $\mu_0 \in \mathbb{R}$, define the test statistic

$$T = \frac{M - \mu_0}{S/\sqrt{n}} \quad (9.2.9)$$

1. If $\mu = \mu_0$, the statistic T has the student t distribution with $n - 1$ degrees of freedom.
2. If $\mu \neq \mu_0$ then T has a non-central t distribution with $n - 1$ degrees of freedom and non-centrality parameter $\frac{\mu - \mu_0}{\sigma/\sqrt{n}}$.

In case (b), the graph of the probability density function of T is much (but not exactly) the same as that of the ordinary t distribution with $n - 1$ degrees of freedom, but shifted to the right or left by the non-centrality parameter, depending on whether $\mu > \mu_0$ or $\mu < \mu_0$.

For $\alpha \in (0, 1)$, each of the following tests has significance level α :

1. Reject $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ if and only if $T < -t_{n-1}(1 - \alpha/2)$ or $T > t_{n-1}(1 - \alpha/2)$ if and only if $M < \mu_0 - t_{n-1}(1 - \alpha/2)\frac{S}{\sqrt{n}}$ or $T > \mu_0 + t_{n-1}(1 - \alpha/2)\frac{S}{\sqrt{n}}$.
2. Reject $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$ if and only if $T > t_{n-1}(1 - \alpha)$ if and only if $M > \mu_0 + t_{n-1}(1 - \alpha)\frac{S}{\sqrt{n}}$.
3. Reject $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0$ if and only if $T < -t_{n-1}(1 - \alpha)$ if and only if $M < \mu_0 - t_{n-1}(1 - \alpha)\frac{S}{\sqrt{n}}$.

Proof

In part (a), T has the chi-square distribution with $n - 1$ degrees of freedom under H_0 . So if H_0 is true, the probability of falsely rejecting H_0 is α by definition of the quantiles. In parts (b) and (c), T has a non-central t distribution with $n - 1$ degrees of freedom under H_0 , as discussed [above](#). Hence if H_0 is true, the maximum type 1 error probability α occurs when $\mu = \mu_0$. The decision rules in terms of M are equivalent to the corresponding ones in terms of T by simple algebra.

Part (a) is the standard two-sided test, while (b) is the right-tailed test and (c) is the left-tailed test. Note that in each case, the hypothesis test is the dual of the corresponding interval estimate constructed in the section on Estimation in the Normal Model.

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

1. $M - t_{n-1}(1 - \alpha/2)\frac{S}{\sqrt{n}} \leq \mu_0 \leq M + t_{n-1}(1 - \alpha/2)\frac{S}{\sqrt{n}}$

2. $\mu_0 \leq M + t_{n-1}(1 - \alpha) \frac{S}{\sqrt{n}}$
3. $\mu_0 \geq M - t_{n-1}(1 - \alpha) \frac{S}{\sqrt{n}}$

Proof

This follows from the [previous result](#). In each case, we start with the inequality that corresponds to *not* rejecting H_0 and then solve for μ_0 .

The two-sided test in (a) corresponds to $\alpha/2$ in each tail of the distribution of the test statistic T , under H_0 . This set is said to be *unbiased*. But of course we can construct other biased tests by partitioning the confidence level α between the left and right tails in a non-symmetric way.

For every $\alpha, p \in (0, 1)$, the following test has significance level α : Reject $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ if and only if $T < t_{n-1}(\alpha - p\alpha)$ or $T \geq t_{n-1}(1 - p\alpha)$ if and only if $M < \mu_0 + t_{n-1}(\alpha - p\alpha) \frac{S}{\sqrt{n}}$ or $M > \mu_0 + t_{n-1}(1 - p\alpha) \frac{S}{\sqrt{n}}$.

1. $p = \frac{1}{2}$ gives the symmetric, unbiased test.
2. $p \downarrow 0$ gives the left-tailed test.
3. $p \uparrow 1$ gives the right-tailed test.

Proof

Once again, H_0 is a simple hypothesis, and under H_0 the test statistic T has the student t distribution with $n - 1$ degrees of freedom. So if H_0 is true, the probability of falsely rejecting H_0 is α by definition of the quantiles. Parts (a)–(c) follow from properties of the quantile function.

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.

The P -values of the standard tests [above](#) are respectively

1. $2[1 - \Phi_{n-1}(|T|)]$
2. $1 - \Phi_{n-1}(T)$
3. $\Phi_{n-1}(T)$

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 $\mu_0 = 1$. Run the experiment 1000 times 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.

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. 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 t tests [above](#) 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 above [two-sided](#), [left-tailed](#), and [right-tailed](#) cases.

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 for the normal pivot variable, in the [two-sided](#) and [one-sided](#) cases.

Tests of the Standard Deviation

For our next discussion, we will construct hypothesis tests for the distribution standard deviation σ . So our assumption is that σ is unknown, and of course almost always, μ would be unknown as well.

For a conjectured value $\sigma_0 \in (0, \infty)$, define the test statistic

$$V = \frac{n-1}{\sigma_0^2} S^2 \quad (9.2.10)$$

1. If $\sigma = \sigma_0$, then V has the chi-square distribution with $n - 1$ degrees of freedom.
2. If $\sigma \neq \sigma_0$ then V has the gamma distribution with shape parameter $(n - 1)/2$ and scale parameter $2\sigma^2/\sigma_0^2$.

Recall that the ordinary chi-square distribution with $n - 1$ degrees of freedom is the gamma distribution with shape parameter $(n - 1)/2$ and scale parameter $\frac{1}{2}$. So in case (b), the ordinary chi-square distribution is scaled by σ^2/σ_0^2 . In particular, the scale factor is greater than 1 if $\sigma > \sigma_0$ and less than 1 if $\sigma < \sigma_0$.

For every $\alpha \in (0, 1)$, the following test has significance level α :

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

Proof

The logic is largely the same as with our other hypothesis test. In part (a), H_0 is a simple hypothesis, and under H_0 , the test statistic V has the chi-square distribution with $n - 1$ degrees of freedom. So if H_0 is true, the probability of falsely rejecting H_0 is α by definition of the quantiles. In parts (b) and (c), V has the more general gamma distribution under H_0 , as discussed [above](#). If H_0 is true, the maximum type 1 error probability is α and occurs when $\sigma = \sigma_0$.

Part (a) is the unbiased, two-sided test that corresponds to $\alpha/2$ in each tail of the chi-square distribution of the test statistic V , under H_0 . Part (b) is the left-tailed test and part (c) is the right-tailed test. Once again, we have a duality between the hypothesis tests and the interval estimates constructed in the section on Estimation in the Normal Model.

For each of the tests in [above](#), we fail to reject H_0 at significance level α if and only if σ_0^2 is in the corresponding $1 - \alpha$ confidence interval. That is

1. $\frac{n-1}{\chi_{n-1}^2(1-\alpha/2)} S^2 \leq \sigma_0^2 \leq \frac{n-1}{\chi_{n-1}^2(\alpha/2)} S^2$
2. $\sigma_0^2 \leq \frac{n-1}{\chi_{n-1}^2(\alpha)} S^2$
3. $\sigma_0^2 \geq \frac{n-1}{\chi_{n-1}^2(1-\alpha)} S^2$

Proof

This follows from the [previous result](#). In each case, we start with the inequality that corresponds to *not* rejecting H_0 and then solve for σ_0^2 .

As before, we can construct more general two-sided tests by partitioning the significance level α between the left and right tails of the chi-square distribution in an arbitrary way.

For every $\alpha, p \in (0, 1)$, the following test has significance level α : Reject $H_0 : \sigma = \sigma_0$ versus $H_1 : \sigma \neq \sigma_0$ if and only if $V \leq \chi_{n-1}^2(\alpha - p\alpha)$ or $V \geq \chi_{n-1}^2(1 - p\alpha)$ if and only if $S^2 < \chi_{n-1}^2(\alpha - p\alpha) \frac{\sigma_0^2}{n-1}$ or $S^2 > \chi_{n-1}^2(1 - p\alpha) \frac{\sigma_0^2}{n-1}$.

1. $p = \frac{1}{2}$ gives the equal-tail test.
2. $p \downarrow 0$ gives the left-tail test.
3. $p \uparrow 1$ gives the right-tail test.

Proof

As before, H_0 is a simple hypothesis, and under H_0 the test statistic V has the chi-square distribution with $n - 1$ degrees of freedom. So if H_0 is true, the probability of falsely rejecting H_0 is α by definition of the quantiles. Parts (a)–(c) follow from properties of the quantile function.

Recall again 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. The power functions of the tests for σ can be expressed in terms of the distribution function G_{n-1} of the chi-square distribution with $n - 1$ degrees of freedom.

The power function of the general two-sided test [above](#) is given by the following formula, and satisfies the given properties:

$$Q(\sigma) = 1 - G_{n-1} \left(\frac{\sigma_0^2}{\sigma^2} \chi_{n-1}^2 (1 - p \alpha) \right) + G_{n-1} \left(\frac{\sigma_0^2}{\sigma^2} \chi_{n-1}^2 (\alpha - p \alpha) \right) \quad (9.2.11)$$

1. Q is decreasing on $(-\infty, \sigma_0)$ and increasing on (σ_0, ∞) .
2. $Q(\sigma_0) = \alpha$.
3. $Q(\sigma) \rightarrow 1$ as $\sigma \uparrow \infty$ and $Q(\sigma) \rightarrow 1$ as $\sigma \downarrow 0$.

The power function of the left-tailed test in [above](#) is given by the following formula, and satisfies the given properties:

$$Q(\sigma) = 1 - G_{n-1} \left(\frac{\sigma_0^2}{\sigma^2} \chi_{n-1}^2 (1 - \alpha) \right) \quad (9.2.12)$$

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

The power function for the right-tailed test [above](#) is given by the following formula, and satisfies the given properties:

$$Q(\sigma) = G_{n-1} \left(\frac{\sigma_0^2}{\sigma^2} \chi_{n-1}^2 (\alpha) \right) \quad (9.2.13)$$

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

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. Record the relative frequency of rejecting the null hypothesis and plot the empirical power curve.

1. Two-sided test
2. Left-tailed test
3. Right-tailed test

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. 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.

1. Two-sided confidence interval
2. Confidence lower bound
3. Confidence upper bound

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^2 are less robust to deviations from the assumption of normality. The following exercises explore these ideas.

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. For each configuration, note the relative frequency of rejecting H_0 . When H_0 is true, compare the relative frequency with the significance level.

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. 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.

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. For each configuration, note the relative frequency of rejecting H_0 . When H_0 is true, compare the relative frequency with the significance level.

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. 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

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 $H_1 : \mu \neq 10$.

1. Suppose that a sample of 100 parts has mean 10.1. Perform the test at the 0.1 level of significance.
2. Compute the P -value for the data in (a).
3. Compute the power of the test in (a) at $\mu = 10.05$.
4. Compute the approximate sample size needed for significance level 0.1 and power 0.8 when $\mu = 10.05$.

Answer

1. Test statistic 3.33, critical values ± 1.645 Reject H_0 .
2. $P = 0.0010$
3. The power of the test at 10.05 is approximately 0.0509.
4. Sample size 223

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:

1. $H_0 : \mu \geq 250$ versus $H_1 : \mu < 250$
2. $H_0 : \sigma \geq 7$ versus $H_1 : \sigma < 7$

Answer

1. Test statistic -3.464 critical value -1.665 Reject H_0 .
2. $P < 0.0001$ so reject H_0 .

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

1. $\mu > 300$?
2. $\sigma > 20$?

Answer

1. Test statistic 2.828, critical value 1.2988. Reject H_0 .
2. $P = 0.0071$ so reject H_0 .

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

1. $\mu > 8$?
2. $\sigma < 1.5$?

Answer

1. Test statistic 2.0, critical value 2.363. Fail to reject H_0 .
2. $P < 0.0001$ so reject H_0 .

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 $\mu < 7.00$?

Answer

Test statistic -1 , critical value -2.328 Fail to reject H_0 .

Data Analysis Exercises

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.

Answer

Test statistic 15.49, critical value 2.6270. Reject H_0 .

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.

Answer

Test statistic -1.269 , critical value -1.7017 Fail to reject H_0 .

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.

Answer

Test statistic -3.730 , critical value ± 1.6749 Reject H_0 .

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

1. The mean petal length of Setosa irises differs from 15 mm.
2. The mean petal length of Verginica irises is greater than 52 mm.
3. The mean petal length of Versicolor irises is less than 44 mm.

Answer

1. Test statistic -1.563 , critical values ± 1.672 Fail to reject H_0 .
2. Test statistic 4.556, critical value 1.2988. Reject H_0 .
3. Test statistic -1.028 , critical value -1.2988 Fail to Reject H_0 .