

9.4: Tests in the Two-Sample Normal Model

In this section, we will study hypothesis tests in the two-sample normal model and in the bivariate normal model. This section parallels the section on Estimation in the Two Sample Normal Model in the chapter on Interval Estimation.

The Two-Sample Normal Model

Suppose that $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a random sample of size m from the normal distribution with mean μ and standard deviation σ , and that $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ is a random sample of size n from the normal distribution with mean ν and standard deviation τ . Moreover, suppose that the samples \mathbf{X} and \mathbf{Y} are independent.

This type of situation arises frequently when the random variables represent a measurement of interest for the objects of the population, and the samples correspond to two different treatments. For example, we might be interested in the blood pressure of a certain population of patients. The \mathbf{X} vector records the blood pressures of a control sample, while the \mathbf{Y} vector records the blood pressures of the sample receiving a new drug. Similarly, we might be interested in the yield of an acre of corn. The \mathbf{X} vector records the yields of a sample receiving one type of fertilizer, while the \mathbf{Y} vector records the yields of a sample receiving a different type of fertilizer.

Usually our interest is in a comparison of the parameters (either the mean or variance) for the two sampling distributions. In this section we will construct tests for the difference of the means and the ratio of the variances. As with previous estimation problems we have studied, the procedures vary depending on what parameters are known or unknown. Also as before, key elements in the construction of the tests are the sample means and sample variances and the special properties of these statistics when the sampling distribution is normal.

We will use the following notation for the sample mean and sample variance of a generic sample $\mathbf{U} = (U_1, U_2, \dots, U_k)$:

$$M(\mathbf{U}) = \frac{1}{k} \sum_{i=1}^k U_i, \quad S^2(\mathbf{U}) = \frac{1}{k-1} \sum_{i=1}^k [U_i - M(\mathbf{U})]^2 \quad (9.4.1)$$

Tests of the Difference in the Means with Known Standard Deviations

Our first discussion concerns tests for the difference in the means $\nu - \mu$ under the assumption that the standard deviations σ and τ are known. This is often, but not always, an unrealistic assumption. In some statistical problems, the variances are stable, and are at least approximately known, while the means may be different because of different treatments. Also this is a good place to start because the analysis is fairly easy.

For a conjectured difference of the means $\delta \in \mathbb{R}$, define the test statistic

$$Z = \frac{[M(\mathbf{Y}) - M(\mathbf{X})] - \delta}{\sqrt{\sigma^2/m + \tau^2/n}} \quad (9.4.2)$$

1. If $\nu - \mu = \delta$ then Z has the standard normal distribution.
2. If $\nu - \mu \neq \delta$ then Z has the normal distribution with mean $[(\nu - \mu) - \delta] / \sqrt{\sigma^2/m + \tau^2/n}$ and variance 1.

Proof

From properties of normal samples, $M(\mathbf{X})$ has a normal distribution with mean μ and variance σ^2/m and similarly $M(\mathbf{Y})$ has a normal distribution with mean ν and variance τ^2/n . Since the samples are independent, $M(\mathbf{X})$ and $M(\mathbf{Y})$ are independent, so $M(\mathbf{Y}) - M(\mathbf{X})$ has a normal distribution with mean $\nu - \mu$ and variance $\sigma^2/m + \tau^2/n$. The final result then follows since Z is a linear function of $M(\mathbf{Y}) - M(\mathbf{X})$.

Of course (b) actually subsumes (a), but we separate them because the two cases play an important role in the hypothesis tests. In part (b), the non-zero mean can be viewed as a *non-centrality parameter*.

As usual, for $p \in (0, 1)$, let $z(p)$ denote the quantile of order p for the standard normal distribution. For selected values of p , $z(p)$ can be obtained from the special distribution calculator or from most statistical software packages. Recall also by symmetry that $z(1-p) = -z(p)$.

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

1. Reject $H_0 : \nu - \mu = \delta$ versus $H_1 : \nu - \mu \neq \delta$ if and only if $Z < -z(1 - \alpha/2)$ or $Z > z(1 - \alpha/2)$ if and only if $M(\mathbf{Y}) - M(\mathbf{X}) > \delta + z(1 - \alpha/2)\sqrt{\sigma^2/m + \tau^2/n}$ or $M(\mathbf{Y}) - M(\mathbf{X}) < \delta - z(1 - \alpha/2)\sqrt{\sigma^2/m + \tau^2/n}$.
2. Reject $H_0 : \nu - \mu \geq \delta$ versus $H_1 : \nu - \mu < \delta$ if and only if $Z < -z(1 - \alpha)$ if and only if $M(\mathbf{Y}) - M(\mathbf{X}) < \delta - z(1 - \alpha)\sqrt{\sigma^2/m + \tau^2/n}$.
3. Reject $H_0 : \nu - \mu \leq \delta$ versus $H_1 : \nu - \mu > \delta$ if and only if $Z > z(1 - \alpha)$ if and only if $M(\mathbf{Y}) - M(\mathbf{X}) > \delta + z(1 - \alpha)\sqrt{\sigma^2/m + \tau^2/n}$.

Proof

This follows the same logic that we have seen before. In part (a), H_0 is a simple hypothesis, and under this hypothesis Z has the standard normal distribution. Thus, if H_0 is true then the probability of falsely rejecting H_0 is α by definition of the quantiles. In parts (b) and (c), H_0 specifies a range of values of $\nu - \mu$, and under H_0 , Z has a nonstandard normal distribution, as described [above](#). But the largest type 1 error probability is α and occurs when $\nu - \mu = \delta$. The decision rules in terms of $M(\mathbf{Y}) - M(\mathbf{X})$ are equivalent to those in terms of Z by simple algebra.

For each of the tests above, we *fail* to reject H_0 at significance level α if and only if δ is in the corresponding $1 - \alpha$ level confidence interval.

1. $[M(\mathbf{Y}) - M(\mathbf{X})] - z(1 - \alpha/2)\sqrt{\sigma^2/m + \tau^2/n} \leq \delta \leq [M(\mathbf{Y}) - M(\mathbf{X})] + z(1 - \alpha/2)\sqrt{\sigma^2/m + \tau^2/n}$
2. $\delta \leq [M(\mathbf{Y}) - M(\mathbf{X})] + z(1 - \alpha)\sqrt{\sigma^2/m + \tau^2/n}$
3. $\delta \geq [M(\mathbf{Y}) - M(\mathbf{X})] - z(1 - \alpha)\sqrt{\sigma^2/m + \tau^2/n}$

Proof

These results follow from the previous results [above](#). In each case, we start with the inequality that corresponds to *not* rejecting the null hypothesis and solve for δ .

Tests of the Difference of the Means with Unknown Standard Deviations

Next we will construct tests for the difference in the means $\nu - \mu$ under the more realistic assumption that the standard deviations σ and τ are unknown. In this case, it is more difficult to find a suitable test statistic, but we can do the analysis in the special case that the standard deviations are the same. Thus, we will assume that $\sigma = \tau$, and the common value σ is unknown. This assumption is reasonable if there is an inherent variability in the measurement variables that does not change even when different treatments are applied to the objects in the population. Recall that the *pooled estimate* of the common variance σ^2 is the weighted average of the sample variances, with the degrees of freedom as the weight factors:

$$S^2(\mathbf{X}, \mathbf{Y}) = \frac{(m-1)S^2(\mathbf{X}) + (n-1)S^2(\mathbf{Y})}{m+n-2} \quad (9.4.3)$$

The statistic $S^2(\mathbf{X}, \mathbf{Y})$ is an unbiased and consistent estimator of the common variance σ^2 .

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

$$T = \frac{[M(\mathbf{Y}) - M(\mathbf{X})] - \delta}{S(\mathbf{X}, \mathbf{Y})\sqrt{1/m + 1/n}} \quad (9.4.4)$$

1. If $\nu - \mu = \delta$ then T has the t distribution with $m + n - 2$ degrees of freedom,
2. If $\nu - \mu \neq \delta$ then T has a non-central t distribution with $m + n - 2$ degrees of freedom and non-centrality parameter

$$\frac{(\nu - \mu) - \delta}{\sigma\sqrt{1/m + 1/n}} \quad (9.4.5)$$

Proof

Part (b) actually subsumes part (a), since the ordinary t distribution is a special case of the non-central t distribution, with non-centrality parameter 0. With some basic algebra, we can write T in the form

$$T = \frac{Z + a}{\sqrt{V/(m+n-2)}} \quad (9.4.6)$$

where Z is the standard score of $M(\mathbf{Y}) - M(\mathbf{X})$, a is the non-centrality parameter given in the theorem, and $V = \frac{m+n-2}{\sigma^2} S^2(\mathbf{X}, \mathbf{Y})$. So Z has the standard normal distribution, V has the chi-square distribution with $m+n-2$ degrees of freedom, and Z and V are independent. Thus by definition, T has the non-central t distribution with $m+n-2$ degrees of freedom and non-centrality parameter a .

As usual, for $k > 0$ and $p \in (0, 1)$, let $t_k(p)$ denote the quantile of order p for the t distribution with k degrees of freedom. For selected values of k and p , values of $t_k(p)$ can be computed from the special distribution calculator, or from most statistical software packages. Recall also that, by symmetry, $t_k(1-p) = -t_k(p)$.

The following tests have significance level α :

1. Reject $H_0 : \nu - \mu = \delta$ versus $H_1 : \nu - \mu \neq \delta$ if and only if $T < -t_{m+n-2}(1-\alpha/2)$ or $T > t_{m+n-2}(1-\alpha/2)$ if and only if $M(\mathbf{Y}) - M(\mathbf{X}) > \delta + t_{m+n-2}(1-\alpha/2)\sqrt{\sigma^2/m + \tau^2/n}$ or $M(\mathbf{Y}) - M(\mathbf{X}) < \delta - t_{m+n-2}(1-\alpha/2)\sqrt{\sigma^2/m + \tau^2/n}$
2. Reject $H_0 : \nu - \mu \geq \delta$ versus $H_1 : \nu - \mu < \delta$ if and only if $T \leq -t_{m+n-2}(1-\alpha)$ if and only if $M(\mathbf{Y}) - M(\mathbf{X}) < \delta - t_{m+n-2}(1-\alpha)\sqrt{\sigma^2/m + \tau^2/n}$
3. Reject $H_0 : \nu - \mu \leq \delta$ versus $H_1 : \nu - \mu > \delta$ if and only if $T \geq t_{m+n-2}(1-\alpha)$ if and only if $M(\mathbf{Y}) - M(\mathbf{X}) > \delta + t_{m+n-2}(1-\alpha)\sqrt{\sigma^2/m + \tau^2/n}$

Proof

This follows the same logic that we have seen before. In part (a), H_0 is a simple hypothesis, and under this hypothesis T has the t distribution with $m+n-2$ degrees of freedom. Thus, if H_0 is true then the probability of falsely rejecting H_0 is α by definition of the quantiles. In parts (b) and (c), H_0 specifies a range of values of $\nu - \mu$, and under H_0 , T has a non-central t distribution, as described [above](#). But the largest type 1 error probability is α and occurs when $\nu - \mu = \delta$. The decision rules in terms of $M(\mathbf{Y}) - M(\mathbf{X})$ are equivalent to those in terms of T by simple algebra.

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

1. $[M(\mathbf{Y}) - M(\mathbf{X})] - t_{m+n-2}(1-\alpha/2)\sqrt{\sigma^2/m + \tau^2/n} \leq \delta \leq [M(\mathbf{Y}) - M(\mathbf{X})] + t_{m+n-2}(1-\alpha/2)\sqrt{\sigma^2/m + \tau^2/n}$
2. $\delta \leq [M(\mathbf{Y}) - M(\mathbf{X})] + t_{m+n-2}(1-\alpha)\sqrt{\sigma^2/m + \tau^2/n}$
3. $\delta \geq [M(\mathbf{Y}) - M(\mathbf{X})] - t_{m+n-2}(1-\alpha)\sqrt{\sigma^2/m + \tau^2/n}$

Proof

These results follow from the previous results [above](#). In each case, we start with the inequality that corresponds to *not* rejecting the null hypothesis and solve for δ .

Tests of the Ratio of the Variances

Next we will construct tests for the ratio of the distribution variances τ^2/σ^2 . So the basic assumption is that the variances, and of course the means μ and ν are unknown.

For a conjectured $\rho \in (0, \infty)$, define the test statistics

$$F = \frac{S^2(\mathbf{X})}{S^2(\mathbf{Y})} \rho \quad (9.4.7)$$

1. If $\tau^2/\sigma^2 = \rho$ then F has the F distribution with $m-1$ degrees of freedom in the numerator and $n-1$ degrees of freedom in the denominator.
2. If $\tau^2/\sigma^2 \neq \rho$ then F has a scaled F distribution with $m-1$ degrees of freedom in the numerator, $n-1$ degrees of freedom in the denominator, and scale factor $\rho \frac{\sigma^2}{\tau^2}$.

Proof

Part (b) actually subsumes part (a) when $\rho = \tau^2/\sigma^2$, so we will just prove (b). Note that

$$F = \left(\frac{S^2(\mathbf{X})/\sigma^2}{S^2(\mathbf{Y})/\tau^2} \right) \rho \frac{\sigma^2}{\tau^2} \quad (9.4.8)$$

But $S^2(\mathbf{X})/\sigma^2$ has the chi-square distribution with $m - 1$ degrees of freedom, $S^2(\mathbf{Y})/\tau^2$ has the chi-square distribution with $n - 1$ degrees of freedom, and the variables are independent. Hence the ratio has the F distribution with $m - 1$ degrees of freedom in the numerator and $n - 1$ degrees of freedom in the denominator

The following tests have significance level α :

1. Reject $H_0 : \tau^2/\sigma^2 = \rho$ versus $H_1 : \tau^2/\sigma^2 \neq \rho$ if and only if $F > f_{m-1, n-1}(1 - \alpha/2)$ or $F < f_{m-1, n-1}(\alpha/2)$.
2. Reject $H_0 : \tau^2/\sigma^2 \leq \rho$ versus $H_1 : \tau^2/\sigma^2 > \rho$ if and only if $F < f_{m-1, n-1}(\alpha)$.
3. Reject $H_0 : \tau^2/\sigma^2 \geq \rho$ versus $H_1 : \tau^2/\sigma^2 < \rho$ if and only if $F > f_{m-1, n-1}(1 - \alpha)$.

Proof

The proof is the usual argument. In part (a), H_0 is a simple hypothesis, and under this hypothesis F has the f distribution with $m - 1$ degrees of freedom in the numerator $n - 1$ degrees of freedom in the denominator. Thus, if H_0 is true then the probability of falsely rejecting H_0 is α by definition of the quantiles. In parts (b) and (c), H_0 specifies a range of values of τ^2/σ^2 , and under H_0 , F has a scaled F distribution, as described [above](#). But the largest type 1 error probability is α and occurs when $\tau^2/\sigma^2 = \rho$.

For each of the tests above, we *fail* to reject H_0 at significance level α if and only if ρ_0 is in the corresponding $1 - \alpha$ level confidence interval.

1. $\frac{S^2(\mathbf{Y})}{S^2(\mathbf{X})} F_{m-1, n-1}(\alpha/2) \leq \rho \leq \frac{S^2(\mathbf{Y})}{S^2(\mathbf{X})} F_{m-1, n-1}(1 - \alpha/2)$
2. $\rho \leq \frac{S^2(\mathbf{Y})}{S^2(\mathbf{X})} F_{m-1, n-1}(\alpha)$
3. $\rho \geq \frac{S^2(\mathbf{Y})}{S^2(\mathbf{X})} F_{m-1, n-1}(1 - \alpha)$

Proof

These results follow from the previous results [above](#). In each case, we start with the inequality that corresponds to *not* rejecting the null hypothesis and solve for ρ .

Tests in the Bivariate Normal Model

In this subsection, we consider a model that is superficially similar to the two-sample normal model, but is actually much simpler. Suppose that

$$((X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)) \quad (9.4.9)$$

is a random sample of size n from the bivariate normal distribution of (X, Y) with $\mathbb{E}(X) = \mu$, $\mathbb{E}(Y) = \nu$, $\text{var}(X) = \sigma^2$, $\text{var}(Y) = \tau^2$, and $\text{cov}(X, Y) = \delta$.

Thus, instead of a *pair of samples*, we have a *sample of pairs*. The fundamental difference is that in this model, variables X and Y are measured on the *same* objects in a sample drawn from the population, while in the previous model, variables X and Y are measured on two distinct samples drawn from the population. The bivariate model arises, for example, in *before and after experiments*, in which a measurement of interest is recorded for a sample of n objects from the population, both before and after a treatment. For example, we could record the blood pressure of a sample of n patients, before and after the administration of a certain drug.

We will use our usual notation for the sample means and variances of $\mathbf{X} = (X_1, X_2, \dots, X_n)$ and $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$. Recall also that the sample covariance of (\mathbf{X}, \mathbf{Y}) is

$$S(\mathbf{X}, \mathbf{Y}) = \frac{1}{n-1} \sum_{i=1}^n [X_i - M(\mathbf{X})][Y_i - M(\mathbf{Y})] \quad (9.4.10)$$

(not to be confused with the pooled estimate of the standard deviation in the two-sample model above).

The sequence of differences $\mathbf{Y} - \mathbf{X} = (Y_1 - X_1, Y_2 - X_2, \dots, Y_n - X_n)$ is a random sample of size n from the distribution of $Y - X$. The sampling distribution is normal with

1. $\mathbb{E}(Y - X) = \nu - \mu$
2. $\text{var}(Y - X) = \sigma^2 + \tau^2 - 2\delta$

The sample mean and variance of the sample of differences are

1. $M(\mathbf{Y} - \mathbf{X}) = M(\mathbf{Y}) - M(\mathbf{X})$
2. $S^2(\mathbf{Y} - \mathbf{X}) = S^2(\mathbf{X}) + S^2(\mathbf{Y}) - 2S(\mathbf{X}, \mathbf{Y})$

The sample of differences $\mathbf{Y} - \mathbf{X}$ fits the normal model for a single variable. The section on Tests in the Normal Model could be used to perform tests for the distribution mean $\nu - \mu$ and the distribution variance $\sigma^2 + \tau^2 - 2\delta$.

Computational Exercises

A new drug is being developed to reduce a certain blood chemical. A sample of 36 patients are given a placebo while a sample of 49 patients are given the drug. The statistics (in mg) are $m_1 = 87$, $s_1 = 4$, $m_2 = 63$, $s_2 = 6$. Test the following at the 10% significance level:

1. $H_0 : \sigma_1 = \sigma_2$ versus $H_1 : \sigma_1 \neq \sigma_2$.
2. $H_0 : \mu_1 \leq \mu_2$ versus $H_1 : \mu_1 > \mu_2$ (assuming that $\sigma_1 = \sigma_2$).
3. Based on (b), is the drug effective?

Answer

1. Test statistic 0.4, critical values 0.585, 1.667. Reject H_0 .
2. Test statistic 1.0, critical values ± 1.6625 Fail to reject H_0 .
3. Probably not

A company claims that an herbal supplement improves intelligence. A sample of 25 persons are given a standard IQ test before and after taking the supplement. The before and after statistics are $m_1 = 105$, $s_1 = 13$, $m_2 = 110$, $s_2 = 17$, $s_{1,2} = 190$. At the 10% significance level, do you believe the company's claim?

Answer

Test statistic 2.8, critical value 1.3184. Reject H_0 .

In Fisher's iris data, consider the petal length variable for the samples of Versicolor and Virginica irises. Test the following at the 10% significance level:

1. $H_0 : \sigma_1 = \sigma_2$ versus $H_1 : \sigma_1 \neq \sigma_2$.
2. $H_0 : \mu_1 \leq \mu_2$ versus $H_1 : \mu_1 > \mu_2$ (assuming that $\sigma_1 = \sigma_2$).

Answer

1. Test statistic 1.1, critical values 0.6227, 1.6072. Fail to reject H_0 .
2. Test statistic -11.4 , critical value -1.6602 Reject H_0 .

A plant has two machines that produce a circular rod whose diameter (in cm) is critical. A sample of 100 rods from the first machine has mean 10.3 and standard deviation 1.2. A sample of 100 rods from the second machine has mean 9.8 and standard deviation 1.6. Test the following hypotheses at the 10% level.

1. $H_0 : \sigma_1 = \sigma_2$ versus $H_1 : \sigma_1 \neq \sigma_2$.
2. $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$ (assuming that $\sigma_1 = \sigma_2$).

Answer

1. Test statistic 0.56, critical values 0.7175, 1.3942. Reject H_0 .
2. Test statistic -4.97 , critical values ± 1.645 Reject H_0 .

This page titled [9.4: Tests in the Two-Sample Normal Model](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist \(Random Services\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.