Elements of a statistical test (I)

- The objective of a statistical test is to test a hypothesis concerning the values of one or more population parameters.
- Any statistical test of hypothesis works in exactly the same way and is composed of the same essential elements:
  1. A null hypothesis H_0: the hypothesis to be tested;
  2. Alternative hypothesis H_1: the hypothesis to be accepted in case H_0 is rejected;
  3. Test statistic: function of the sample measurements
  4. Rejection region (RR): specify the values of the test statistic for which the null hypothesis is rejected.

Definition
A type I error is made if H_0 is rejected when H_0 is true. The probability of a type I error is denoted by α.

A type II error is made if H_0 is accepted when H_1 is true. The probability of a type II error is denoted by β.

Elements of a statistical test (II)

	H_0 is true	H_1 is true
H_0 not rejected	Right	Type II error
H_0 is rejected	Type I error	Right

$$\alpha = P(\text{rejecting } H_0 \text{ when } H_0 \text{ is true})$$

$$\beta = P(\text{accepting } H_0 \text{ when } H_1 \text{ is true}).$$

	H_0 is true	H_1 is true
H_0 not rejected	$1-\alpha$	β
H_0 is rejected	α	$1-\beta$

COMMON LARGE-SAMPLE TESTS (I)

- We want to test a hypothesis concerning a parameter θ, based on a random sample $X_1, \cdots, X_n$.
- We will develop a procedure based on an estimator $\widehat{\theta}$ which has an (approximately) normal distribution with a mean of θ and a variance of $\sigma^2_{\widehat{\theta}}$.
- We may wish to test $H_0 : \theta = \theta_0$, versus $H_1 : \theta > \theta_0$.
- If $\widehat{\theta}$ is close to θ_0, it seems reasonable to accept H_0. However, if $\theta > \theta_0$, it is more likely that $\widehat{\theta}$ will be large.
- Then, here the test statistic is $\widehat{\theta}$ and the rejection region is $RR = \{\widehat{\theta} > k\}$ for some choice of k.

Common large-sample tests (II)

- The value of k is determined by fixing the type I error probability α.
- If H_0 is true, then $\widehat{\theta} \sim \mathcal{N}(\theta_0, \sigma^2_{\widehat{\theta}})$, therefore

$$\alpha = P(\widehat{\theta} > k) = P\left(\frac{\widehat{\theta} - \theta_0}{\sigma_{\widehat{\theta}}} > \frac{k - \theta_0}{\sigma_{\widehat{\theta}}}\right)$$

- The appropriate choice for k is $(k - \theta_0)/\sigma_{\widehat{\theta}} = z_\alpha$ with $P(Z > z_\alpha) = \alpha$, that is

$$k = \theta_0 + z_\alpha \sigma_{\widehat{\theta}}.$$

Common large-sample tests (III)

- A test of $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$ would be carried out in an analogous manner, except that $RR = \{z < -z_\alpha\}$ for the statistic $Z = (\widehat{\theta} - \theta_0)/\sigma_{\widehat{\theta}}$.

- If we wish to test $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$, the rejection region is $RR = \{|z| > z_{\alpha/2}\}$.

Calculating type-II error

- Calculating β can be very difficult for some statistical tests, but it is easy for the previous test.
- For the test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$ it is possible to calculate the type II error probability for only specific points in H_1.
- Suppose that the experimenter has a specific alternative say, $\theta = \theta_a$ $(\theta_a > \theta_0)$ in mind. Because $RR = \{\widehat{\theta} : \ \widehat{\theta} > k\}$, we have

$$\begin{aligned}\beta &= P(\widehat{\theta} \text{ not in RR when } H_1 \text{ is true}) \\ &= P(\widehat{\theta} \leq k \text{ when } \theta = \theta_a) \\ &= P\left(\frac{\widehat{\theta} - \theta_a}{\sigma_{\widehat{\theta}}} \leq \frac{k - \theta_a}{\sigma_{\widehat{\theta}}}\right).\end{aligned}$$

- If θ_a is the true value of θ, then $(\widehat{\theta} - \theta_a)/\sigma_{\widehat{\theta}} \sim \mathcal{N}(0, 1)$ and β can be determined by the standard method.

Sample size for the Z test

- For a fixed sample size n, the size of β will depend upon the distance between θ_a and θ_0;
- If θ_a is close to θ_0, β will tend to be large;
- If θ_a is far from θ_0, β will be considerable smaller.
- A large value of β tells us that the sample size is too small.
- Suppose we want to test $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$. If you specify α and β the test depends upon n and k.
- We have

$$\alpha = P\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > \frac{k - \mu_0}{\sigma/\sqrt{n}}\right) = P(Z > z_\alpha)$$

$$\beta = P\left(\frac{\overline{X} - \mu_a}{\sigma/\sqrt{n}} \leq \frac{k - \mu_a}{\sigma/\sqrt{n}}\right) = P(Z \leq -z_\beta)$$

- We obtain $\frac{k-\mu_0}{\sigma/\sqrt{n}} = z_\alpha, \quad \frac{k-\mu_a}{\sigma/\sqrt{n}} = -z_\beta.$
- Eliminating k gives

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

RESULTS OF A TEST: p-VALUE (I)

- The probability α, called the **significance level** is somewhat arbitrary. One experimenter might choose $\alpha = 0.05$, whereas another experimenter may prefer $\alpha = 0.01$
- It is possible that two persons could analyse the same data, one concluding that H_0 should be rejected at the $\alpha = 0.05$ significance level, and the other deciding that H_0 can't be rejected at $\alpha = 0.01$

Definition
If W is a test statistic, the p**-value** is the smallest level of significance, α, for which the observed data indicates that the null-hypothesis should be rejected.

RESULTS OF A TEST: p-VALUE (II)

If an experimenter chooses $\alpha \geq p$-value, H_0 is rejected. Otherwise, if $\alpha < p$-value, H_0 cannot be rejected.

The general method of computing p-values is the following:

- If one were to reject H_0 in favour of H_1 for small values of a test statistic W, the p-value associated with an observed value of W, say w_0 is given by

$$p\text{-value} = P(W \leq w_0, \text{ when } H_0 \text{ is true}).$$

- Analogously, if H_0 should be rejected for large values of W

$$p\text{-value} = P(W \geq w_0, \text{ when } H_0 \text{ is true}).$$

Test and Student's distribution (I)

- We assume that $X_1, \cdots, X_n$ denotes a random sample of size n from a **normal distribution** with unknown mean μ and unknown variance σ^2.
- If $\overline{X}$ and S denote the sample mean and standard deviation respectively, and if $H_0 : \mu = \mu_0$ is true, then
$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1).$$
- Because the t-distribution is symmetric it is clear that the rejection region for a small sample test of H_0 would be determined exactly the same way as for the large-sample Z test.

$$H_0 : \mu = \mu_0,\ H_1 : \begin{cases} \mu > \mu_0 \\ \mu < \mu_0 \\ \mu \neq \mu_0 \end{cases} \quad T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}},\ RR = \begin{cases} t > t_\alpha \\ t < -t_\alpha \\ |t| > t_{\alpha/2} \end{cases}$$

TEST AND STUDENT'S DISTRIBUTION (II)

- A second application of the t-distribution is its use in constructing a small-sample test to compare the means of two normal population that possess equal variances.
- We have independent random samples from each population
$$X_{11}, \cdots, X_{1n_1} \sim \mathcal{N}(\mu_1, \sigma^2),\ X_{21}, \cdots, X_{2n_2} \sim \mathcal{N}(\mu_2, \sigma^2).$$
- Assume that $\overline{X}_i, S_i^2$, $i = 1, 2$ are the corresponding sample means and variances. Then, we showed in Section 8.8 that if
$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$
is the pooled estimator for σ^2, then
$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2).$$
- If $H_0; \mu_1 - \mu_2 = D_0$ for some value of D_0, it follows that if H_0 is true then
$$T = \frac{(\overline{X}_1 - \overline{X}_2) - D_0}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2).$$

Test and Student's distribution (III)

Assumptions: Independent samples from normal distributions with the **same variance**

$$H_0 : \mu_1 - \mu_2 = D_0, \quad H_1 : \begin{cases} \mu_1 - \mu_2 > D_0 \\ \mu_1 - \mu_2 < D_0 \\ \mu_1 - \mu_2 \neq D_0 \end{cases}$$

Test-Statistic:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - D_0}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

Rejection Region: $RR : \begin{cases} t > t_{\alpha, n_1+n_2-2} \\ t < -t_{\alpha, n_1+n_2-2} \\ |t| > t_{\alpha/2, n_1+n_2-2} \end{cases}$

Like the t-test for a single mean, the t-test for comparing two populations means is robust relative to the assumption of normality. It is also robust relative to assumption that $\sigma_1^2 = \sigma_2^2$ when $n_1 = n_2$ are equal (or nearly equal).

Testing hypotheses concerning variances(I)

- We again assume that we have a random sample $X_1, \cdots, X_n$ from a normal distribution with an unknown mean of μ and an unknown variance of σ^2.
- We test $H_0 : \sigma^2 = \sigma_0^2$ for some fixed value σ_0^2 versus various alternative hypotheses.
- We know that $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ has a χ^2 distribution with $n-1$ degrees of freedom when H_0 is true.
- If we desire to test H_0 against $H_1 : \sigma^2 > \sigma_0^2$, we can use this statistic χ^2 has our test statistic, but how should we select the rejection region, RR?
- If H_1 is true, then we would expect S^2 to be larger than σ_0^2. The larger S^2 is relative to σ_0^2, the stronger will be the evidence to support H_1. Thus we see that a rejection region of the form $RR = \{\chi^2 > k\}$ for some constant k would be appropriate.
- If we desire a test for which the probability of a type I error is α, then we use the rejection region $RR = \{\chi^2 > \chi_\alpha^2\}$, where $P(\chi^2 > \chi_\alpha^2) = \alpha$.

Testing hypotheses concerning variances (II)

Assumptions: $X_1, \cdots, X_n$ is a random sample from a normal distribution with mean μ and variance σ^2

$$H_0 : \sigma^2 = \sigma_0^2,$$

$$H_1 : \begin{cases} \sigma^2 > \sigma_0^2 \\ \sigma^2 < \sigma_0^2 \\ \sigma^2 \neq \sigma_0^2 \end{cases}$$

Test Statistic: $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$

Rejection Region: RR: $\begin{cases} \chi^2 > \chi^2_{\alpha, n-1} \\ \chi^2 < \chi^2_{1-\alpha, n-1} \\ \chi^2 > \chi^2_{\alpha/2, n-1} \text{ or } \chi^2 < \chi^2_{1-\alpha/2, n-1} \end{cases}$

Note that χ^2_α is chosen so that $P(\chi^2 > \chi^2_\alpha) = \alpha$ (See Table 6).

Testing hypotheses concerning variances (III)

- Sometimes, we wish to compare the variances if two normal distributions by testing to determine whether or not they are equal.
- For example, suppose that $X_{11}, \cdots, X_{1n_1}$, and $X_{21}, \cdots, X_{2n_2}$ are independent random samples from normal distributions with unknown means and that $V(X_{1i}) = \sigma_1^2$ and $V(X_{2i}) = \sigma_2^2$, where σ_1^2 and σ_2^2 are unknown.
- Suppose we want to test the null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ against the alternative $H_1 : \sigma_1^2 > \sigma_2^2$.
- Because the sample variances S_1^2 and S_2^2 estimate the respective population variances, σ_1^2 and σ_2^2, we would reject H_0 in favour of H_1 if S_1^2 is much larger than S_2^2.
- We use a rejection region of the form
$$RR = \left\{ \frac{S_1^2}{S_2^2} > k \right\},$$
where k is chosen so that the probability of the type I error is α.

Testing hypotheses concerning variances(IV)

- The appropriate value of k depends upon the distribution of S_1^2/S_2^2. Note that $(n_1-1)S_1^2/\sigma_1^2$ and $(n_2-1)S_2^2/\sigma_2^2$ are independent chi-square random variables and it follows that
$$F = \frac{(n_1-1)S_1^2}{\sigma_1^2(n_1-1)} \Big/ \frac{(n_2-1)S_2^2}{\sigma_2^2(n_2-1)} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \sim F(n_1-1, n_2-1).$$
- Under the null hypothesis, $\sigma_1^2 = \sigma_2^2$, then $F = S_1^2/S_2^2$ and the rejection region RR given earlier is equivalent to
$$RR = \{F > k\} = \{F > F_\alpha\},$$
where F_α is the value of the F distribution with $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ such that $P(F > F_\alpha) = \alpha$ (See Table 7).
- Suppose that our alternative hypothesis was $H_1 : \sigma_1^2 < \sigma_2^2$. How would we proceed? We are free to identify either population as population 1. Therefore, if we simply interchange the arbitrary labels of 1 and 2 on the two populations (and the corresponding identifiers on sample sizes, sample variances, etc.), our alternative hypothesis becomes $H_1 : \sigma_1^2 > \sigma_2^2$.

Testing hypotheses concerning variances (V)

Assumptions: Independent samples from normal populations

$$H_0 : \sigma_1^2 = \sigma_2^2, \quad H_1 : \sigma_1^2 > \sigma_2^2$$

Test -Statistic:

$$F = \frac{S_1^2}{S_2^2}$$

Rejection Region:

$$RR = \{F > F_\alpha\},$$

where F_α is chosen so that $P(F > F_\alpha) = \alpha$, when F has $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ degrees of freedom (See Table 7).

Testing hypotheses concerning variances(VI)

- If we wish to test $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_1 : \sigma_1^2 \neq \sigma_2^2$, with type I error α, we could employ $F = S_1^2/S_2^2$ as a test statistic and reject H_0 if the calculated F is in either the upper or lower $\alpha/2$ tail of the F distribution.
- The upper-tail critical values can be determined directly from a table.
- In order to determine the lower-tail critical values, we note that F and $F^{-1} = S_2^2/S_1^2$ both have F distributions, but that the degrees of freedom are interchanged.
- Let F_b^a denote a random variable with an F distribution with $\nu_1 = a$ and $\nu_2 = b$ degrees of freedom, and let $F_{b,\alpha/2}^a$ be such that
$$P(F_b^a > F_{b,\alpha/2}^a) = \alpha/2$$
Then
$$P\left((F_b^a)^{-1} < (F_{b,\alpha/2}^a)^{-1}\right) = \alpha/2$$
and therefore
$$P(F_a^b < (F_{b,\alpha/2}^a)^{-1}) = \alpha/2$$

Testing hypotheses concerning variances (VII)

- That is, the value that cuts off the lower-tail area of $\alpha/2$ for an F_a^b distribution can be found by inverting $F_{b,\alpha/2}^a$. Thus if we use $F = S_1^2/S_2^2$ as a test statistic for testing $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_1 : \sigma_1^2 \neq \sigma_2^2$, the appropriate rejection region is

$$RR : \{F > F_{n_2-1,\alpha/2}^{n_1-1} \text{ or } F < (F_{n_1-1,\alpha/2}^{n_2-1})^{-1}\}.$$

- An equivalent test is obtained as follows. Let n_L and n_S denote the sample sizes associated with the larger and smaller samples variances respectively. Place the large sample variance in the numerator and the smaller sample variance in the denominator of the F statistic and reject $H_0 : \sigma_1^2 = \sigma_2^2$ in favour of $H_1 : \sigma_1^2 \neq \sigma_2^2$ if $F > F_{\alpha/2}$ where $F_{\alpha/2}$ is determined for $\nu_1 = n_L - 1$ and $\nu_2 = n_S - 1$.