3.7 : The Student's t Test

In Section 3.1 we introduced the notion of a hypothesis test, and gave some important definitions, by considering a hypothesis test to determine the mean, μ, of a normal pdf with known variance, $\sigma^2 = 1$, based on a single]sampled value, x. Effectively we constructed the test statistic

$$z = \frac{x - \mu}{\sigma}$$

where $\mu = -2$ under H_1 and $\mu = 2$ under H_2 and z was a RV drawn from a standard normal pdf, $N(0,1)$, with mean zero and unit variance. (N.B. recall that a statistic cannot depend on any *unknown* parameters, but here σ is assumed known and μ is specified exactly under either H_1 and H_2, so it makes sense to regard z as a statistic).

The more realistic situation, on the other hand, is where σ is *not* known *a priori*. In this case we can infer nothing about μ from a single observation since we have no idea of how 'broad' the pdf is. If $n \geq 2$, however, then we can construct a hypothesis test for the value of the true mean, μ, by first determining the sample mean and variance of our random sample.

Suppose we want to test the hypothesis that the true mean takes some specific value, μ_0, i.e. we take as our null and alternative hypotheses:-

$$\text{NH} : \mu = \mu_0 \qquad \text{AH} : \mu \neq \mu_0$$

We construct the following test statistic

$$t = \frac{\hat{\mu} - \mu_0}{\sigma_{\hat{\mu}}}$$

where $\hat{\mu}$ is the sample mean, i.e. $\hat{\mu} = \frac{1}{n} \sum x_i$, and $\sigma_{\hat{\mu}}$ is the **standard error on the mean** (see handout), i.e.

$$\sigma_{\hat{\mu}} = \left[\frac{1}{n(n-1)} \sum_{i=1}^{n} (x_i - \hat{\mu})^2 \right]^{\frac{1}{2}}$$

Thus

$$t = \frac{\sqrt{n(n-1)}(\hat{\mu} - \mu_0)}{[\sum (x_i - \hat{\mu})^2]^{\frac{1}{2}}}$$

t has a pdf known as the student's t distribution. It is similar in shape to a standard normal pdf, $N(0,1)$, but with wider 'wings' (i.e. positive kurtosis) and its shape also depends on n – see the figure in the statistical tables. The pdf of t has $\nu = n - 1$ degrees of freedom. Thus, for a sample of size n, to carry out our hypothesis test we determine the value of the student's t statistic under the NH that $\mu = \mu_0$ and compare this value with the critical values of the pdf, for the appropriate number of degrees of freedom.

Note that the hypothesis test given above, where our AH is $\mu \neq \mu_0$, calls for a **two-tailed test**, since a value of t significantly larger *or* smaller than zero would argue in favour of the AH. If, on the other hand, we want only to test if $\mu > \mu_0$ ($\mu < \mu_0$) then we would take as our critical region the upper (lower) tail of the student's t distribution.

3.8 : Difference of Means

Let $\{x_1, ..., x_{n_1}\}$ and $\{y_1, ..., y_{n_2}\}$ be iid random samples from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively, where $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Suppose we wish to test the NH that $\mu_1 - \mu_2 = \mu_0$, i.e. the means of the pdfs from which the two samples are drawn differ by a fixed amount. Under the NH then the difference of the sample means, $\hat{\mu_1} - \hat{\mu_2}$, is a normal pdf with mean μ_0 and variance $\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$. If σ^2 were known then the appropriate test statistic to test the NH would be

$$z = \frac{(\hat{\mu_1} - \hat{\mu_2}) - \mu_0}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

which under the NH would have pdf $N(0,1)$. If, as in Section 3.7, σ^2 is *not* known *a priori*, then we use the test statistic

$$t = \frac{(\hat{\mu_1} - \hat{\mu_2}) - \mu_0}{\hat{\sigma} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where

$$\hat{\sigma} = \left[\frac{1}{n_1 + n_2 - 2} \left(\sum_{i=1}^{n_1} (x_i - \hat{\mu_1})^2 + \sum_{i=1}^{n_2} (y_i - \hat{\mu_2})^2 \right) \right]^{\frac{1}{2}}$$

(i.e. $\hat{\sigma}^2$ is the weighted mean of the unbiased estimators, from the first and second samples, of the variance on a single observation – see lectures)

Under the NH t has the student's t distribution with $\nu = n_1 + n_2 - 2$ degrees of freedom. Clearly to test if $\mu_1 = \mu_2$ we simply set $\mu_0 = 0$.

3.9 : F Test for the Ratio of Variances

Let $\{x_1, ..., x_{n_1}\}$ and $\{y_1, ..., y_{n_2}\}$ be iid random samples from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively. Unlike Section 3.8, we do **not** now assume that $\sigma_1^2 = \sigma_2^2$. In fact we specifically want a simple hypothesis test of:-

$$\text{NH}: \quad \sigma_1^2 = \sigma_2^2 \qquad \text{AH}: \quad \sigma_1^2 \neq \sigma_2^2$$

We use the test statistic, f, defined by:-

$$f = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_1^2}$$

where

$$\hat{\sigma}_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(x_i - \hat{\mu_1})^2 \quad \text{and} \quad \hat{\sigma}_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(y_i - \hat{\mu_2})^2$$

Under the NH, that the two distributions have equal variance, this test statistic has a pdf known as the F distribution, with $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ degrees of freedom. (We usually write this as F_{ν_1,ν_2}). The pdf of F has a complicated analytic form which need not concern us here. The essential idea of how it is used in practice is that, if the NH is true, then values of the f statistic significantly larger or smaller than unity are unlikely. Hence, by computing the f statistic and comparing the observed value with tabulated critical values, we can make a decision whether to accept or reject the NH that the pdfs from which the two samples were drawn have equal variance.

Typical astrophysical problems to which the F test can be applied include comparing the luminosity function of stars of different spectral types, or galaxies of different morphological type. One can also test, for example, whether the spatial distribution of galaxies of different morphological types is significantly different – e.g. do spirals and ellipticals have the same spatial distribution in galaxy clusters, or are ellipticals found preferentially in the cores of clusters. (Primordial spirals which originally formed in cluster cores are thought to have been torn apart and 'cannibalised' by ellipticals because of the strong tidal forces in the cluster core, so that they are not found in the cores of clusters today). See example sheet 4 for some similar applications of the F test.

3.10 : Hypothesis Tests on the Sample Correlation Coefficient

The final type of hypothesis test which we consider is associated with testing whether two variables are statistically independent, which we can do by considering the value of the **sample correlation coefficient**. In Section 1.9 we defined the covariance of two RVs, X and Y, as

$$\text{cov}(X,Y) = E[(X - \mu_\text{x})(Y - \mu_\text{y})]$$

and the correlation coefficient, ρ, as

$$\rho = \frac{\text{cov}(X,Y)}{\sigma_\text{x}\sigma_\text{y}}$$

While we defined ρ in Section 1.9 in the context of its role as a parameter of the bivariate normal distribution, one can define the covariance and correlation coefficient of *any* two RVs (i.e. with *any* bivariate distribution) using the above formulae. As in the case of a bivariate normal pdf, it follows that

$$X \text{ and } Y \text{ are independent} \quad \leftrightarrow \quad \text{cov}(X,Y) = 0 \quad \leftrightarrow \quad \rho = 0$$

We estimate ρ by the **sample correlation coefficient**, $\hat{\rho}$, defined by:-

$$\hat{\rho} \quad = \quad \frac{\sum(x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y)}{\sqrt{[\sum(x_i - \hat{\mu}_x)^2]\,[\sum(y_i - \hat{\mu}_y)^2]}}$$

where, as usual, $\hat{\mu}_x$ and $\hat{\mu}_y$ denote the sample means of X and Y respectively, and all sums are over $1, \ldots, n$, for sample size, n. $\hat{\rho}$ is also often denoted by r, and is referred to as 'Pearson's correlation coefficient'.

Of course, if X and Y *do* have a bivariate normal pdf, then ρ corresponds precisely to the parameter defined in Section 1.9. To test hypotheses about ρ we need to know the sampling distribution of $\hat{\rho}$. We consider two special cases, both of which are when X and Y have a bivariate normal pdf.

3.10.1 : $\rho = 0$ (i.e. X and Y are independent)

If $\rho = 0$, then the statistic

$$t \quad = \quad \frac{\hat{\rho}\sqrt{n-2}}{1-\hat{\rho}^2}$$

has a student's t distribution, with $\nu = n - 2$ degrees of freedom. Hence, we can use t to test the hypothesis that X and Y are independent. (See example sheets and lectures).

3.10.2 : $\rho = \rho_0 \neq 0$

In this case, then **for large samples**, the statistic

$$z \quad = \quad \frac{1}{2}\log_e\left(\frac{1+\hat{\rho}}{1-\hat{\rho}}\right)$$

has an approximately normal pdf with mean, μ_z and variance σ_z^2 given by

$$\mu_z = \frac{1}{2}\log_e\left(\frac{1+\rho_0}{1-\rho_0}\right) \qquad \sigma_z^2 = \frac{1}{n-3}$$