1. Comparing a single patient versus a group of controls (and SPM)

1.1 Z-scores versus one-sample T-tests

If we knew the mean, μ, and standard deviation, σ, of the control "population", we could construct the "Z-score":

$$Z = \frac{\mu - p}{\sigma}$$

where *p* is the single patient score, and determine the probability of obtaining a Z-score this large from the "Z-distribution" (or "normal" or "Gaussian" distribution). In practice, Z-scores are often tested this way when the group of controls is large, such that one has very reliable estimates of the population mean and standard deviation (e.g, when using "norm-based" tests).

However, when we estimate μ and σ from a small sample of the population, we need to acknowledge that our values of μ and σ are estimated quantities that have associated errors.

Firstly, we need to take into account the distribution of the mean (i.e, the range of values of the mean that would obtain if we took many samples from the population). Given a sample of *N* (control) scores, c_i, the mean is estimated as:

$$\hat{\mu} = \frac{\sum_i^N c_i}{N}$$

By the "central limit theorem", the distribution of means estimated this way is approximately Gaussian, with an associated standard deviation (called the "standard error of the mean") of:

$$\sigma_\mu = \sigma / \sqrt{N}$$

Secondly, we need to estimate σ. The standard (unbiased) estimate is:

$$\hat{\sigma}^2 = \frac{\sum (c_i - \hat{\mu})^2}{N-1} \qquad [1]$$

However, the distribution of $\hat{\sigma}$'s estimated this way is not Gaussian, but has a positive, skewed distribution. This means we cannot test against Z-distributions anymore. Instead, we construct a "T-score":

$$T = \frac{\hat{\mu} - p}{\hat{\sigma}_\mu}$$

This has a known "T-distribution" ("Student's T"), which is a function of the degrees of freedom (v). Rearranging from the above:

$$T(v) = \frac{\hat{\mu} - p}{\hat{\sigma}/\sqrt{N}} \qquad [2]$$

Because we have used one degree of freedom in estimating σ (i.e, by needing to first estimate μ), the degrees of freedom for the T-test are $v=N-1$. Note that, when v is large, the T-distribution approaches a Z-distribution, so one could use Z-tests instead.

In SPM, such a T-test is equivalent to creating a contrast image of each control subject image minus the patient image, and entering the resulting difference images into a `"one-sample"` T-test (Figure 1, left).

However, this one-sample T-test approach does not assume any variance associated with the patient's score p. If the patient were actually drawn from a population of like patients, we could use information about the variance of that population. This would allow a two-sample T-test.

1.2 Two-sample T-tests

When we have two, independent samples, we can construct a T-statistic to test whether the means of the corresponding populations differ. As with any T-test, we assume that the standard deviation of the two populations is equal. However, we can estimate this common variance in two ways: using either an "unpooled" or a "pooled" estimate. For an unpooled estimate (dropping v for the simplicity):

$$T = \frac{\hat{\mu}_c - \hat{\mu}_p}{\sqrt{\hat{\sigma}_c^2 / N_c + \hat{\sigma}_p^2 / N_p}} \qquad [3]$$

where the c and p subscripts denote controls and patients respectively. For a pooled estimate on the other hand, the T-value is defined as:

$$T = \frac{\hat{\mu}_c - \hat{\mu}_p}{\sqrt{\hat{\sigma}^2 (1/N_c + 1/N_p)}} \qquad \hat{\sigma}^2 = \frac{(N_c - 1)\hat{\sigma}_c^2 + (N_p - 1)\hat{\sigma}_p^2}{N_c + N_p - 2} \qquad [4]$$

In the pooled case, the common variance is estimated by weighting each sample variance by the number of observations in that sample, i.e giving more weight to larger samples (since they are better estimates). When $N_c = N_p$, Equations [3] and [4] are equivalent of course.

Now when the patient group contains only a single patient, $N_p = 1$, then we cannot really estimate the standard deviation $\hat{\sigma}_p$ (from Equation [1]). If however we assume $\hat{\sigma}_p = 0$ (and $\hat{\mu}_p = p$), then Equation [3] reduces to Equation [2], i.e, the unpooled two-sample T-test reduces to one-sample T-test approach.

On the hand, the pooled approach reduces to:

$$\hat{\sigma}^2 = \hat{\sigma}_c^2 \qquad T = \frac{\hat{\mu}_c - p}{\sqrt{\hat{\sigma}_c^2 (1/N_c + 1)}} = \frac{\hat{\mu} - p}{\hat{\sigma}\sqrt{(1/N + 1)}} \qquad [5]$$

(dropping the c subscript). This is sometimes called "Crawford's modified T-test".

Comparing Equation [5] with Equation [2], one can see that the one-sample T-test and pooled two-sample T-test can give quite different answers (the latter being much more conservative)!

Note that SPM always effectively uses a pooled estimate (of the error variance). Thus in SPM, the above pooled two-sample t-test is equivalent to creating an `"unpaired t-test"` and entering the control images for the first group, the single patient image for the second group, and testing the T-contrast [1 -1] (Figure 1, right).
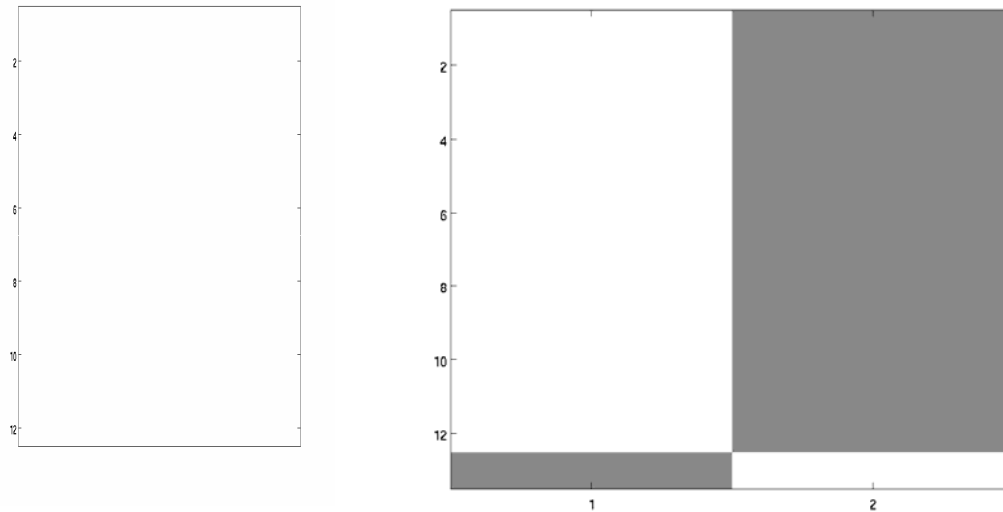
Figure 1. Design Matrices for One- and Two-sample T-tests in SPM ($N_c=12$)

1.3 Which test to use?

For small sample sizes (i.e, small numbers of controls), one must use a T-test rather than Z-test. But should one use a one- or two-sample T-test? It depends on the question one is asking. If the patient is regarded as a unique, single-case (and one assumes the measurement error associated with their score is negligible), then one might prefer the one-sample T-test, in which the patient score is treated as a fixed effect. If however one regards the patient as one of a population of like patients (even if that population is hypothetical, rather than real - i.e, even if no similar such patients have *yet* been found), then one might prefer the two-sample T-test (using a pooled error). Indeed, the pooled two-sample T-test (Equation [5]) is often recommended by statisticians.

2. Comparing a small number of patients versus controls

If you have a large number of controls, you could:

1. Calculate a Z-score for each patient (Section 1.1) and report whether each patient is reliably different using a Z-test.

2. Calculate a Z-score for each patient (Section 1.1) and test the cumulative Z-score, Z_c:

$$Z_c = \frac{\sum_i^{N_p} Z_i}{\sqrt{N}}$$

If you have a small number of controls however, you should probably use a pooled two-sample T-test (Section 1.2):

$$T = \frac{\hat{\mu}_c - \hat{\mu}_p}{\sqrt{\hat{\sigma}^2 (1/N_c + 1/N_p)}} \qquad \hat{\sigma}^2 = \frac{(N_c - 1)\hat{\sigma}_c^2 + (N_p - 1)\hat{\sigma}_p^2}{N_c + N_p - 2}$$