Hypothesis Tests (8 pages; 16/4/16)

(excluding Correlation, χ^2 test for independence or Goodness of Fit, and Analysis of Variance)

[Unless indicated otherwise, a 'large' sample means $n \geq 30$]

General

(i) When accepting/rejecting H_0, it is recommended to express the conclusion both (a) formally and (b) 'in context', in layman's terms.

For example:

"As 13.2 [the test statistic] $>$ 12.7 [the critical value], reject H_0 at the 5% significance level. There is enough evidence to conclude that average rainfall has increased."

(ii) It is also possible to use a confidence interval to test a hypothesis (ie reject H_0 if the value proposed by H_0 lies outside the appropriate interval).

(A) Single samples

(1) Test for mean of a Normal distribution with known variance

$H_0: X \sim N(\mu, \sigma^2) \Rightarrow \overline{X} \sim N(\mu, \frac{\sigma^2}{n})$

Test statistic: $z = \frac{\bar{x}-\mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$ (compare with the critical value from the Normal table).

(2) Test for mean of a Normal distribution with unknown variance, where the sample is large

As the sample is large (usually taken to be ≥ 30), s (based on a divisor of $n-1$) can be assumed to be a reasonably good approximation to σ.

The test statistic is then $z = \frac{\bar{x}-\mu}{\left(\frac{s}{\sqrt{n}}\right)}$

(3) Test for mean of an unknown distribution (with unknown variance), where the sample is large

As the sample size is large, the Central Limit theorem says that $\bar{X}$ approx. $\sim N(\mu, \frac{\sigma^2}{n})$ [$n \geq 30$ also applies here] and s can be assumed to be a reasonably good approximation to σ (again, as the sample size is large).

As in (2), the test statistic is $z = \frac{\bar{x}-\mu}{\left(\frac{s}{\sqrt{n}}\right)}$

(4) Test for mean of a Normal distribution with unknown variance, where the sample is small

To reflect the greater uncertainly caused by approximating σ by s when the sample is small, the t-distribution is used, with

$\upsilon = n - 1$ degrees of freedom.

The test statistic is $t_{n-1} = \frac{\bar{x}-\mu}{\left(\frac{s}{\sqrt{n}}\right)}$ (compare with the critical value from the t table).

(Note that the underlying distribution has to be Normal, in order for the t-distribution to apply. As the sample size increases, the t-value tends to the z-value.)

(5) Test for mean or median of a symmetrical but otherwise unknown distribution, where the sample is small

The Wilcoxon signed-rank test can be used.

This is a test for the median, but can be used approximately for the mean.

Procedure:

(i) H_0: median is M

(ii) Given a sample of x_i of size n , calculate the differences $x_i - M$

(iii) Rank the $x_i - M$ by absolute size, with a rank of 1 for the smallest value of $|x_i - M|$

(iv) Calculate the sum of ranks for the positive differences, and also for the negative differences. The test statistic, W is then the smaller of these sums.

(v) The (lower tail) critical value is obtained from the Wilcoxon signed-rank table. Reject H_0 if W < critical value [ie W is suspiciously small if H_0 is assumed]

[Note: For larger n,

$W \sim$ approx. $N(\frac{1}{4}n(n+1), \frac{1}{24}n(n+1)(2n+1))$, with a continuity correction being applied (presumably); however, (3) may also be applied for large n]

(6) Test for mean or median of an unknown distribution (which cannot be assumed to be symmetrical), where the sample is small

The Sign test can be used.

H_0: median is M

Test statistic is P: the number of positive values of $x_i - M$

$H_0 \Rightarrow P \sim B(n, 0.5)$,

so that the critical value is obtained from the Binomial table

(7) Test for Binomial proportion, for a large sample (using a Normal approximation)

$H_0: X \sim B(n, p)$. If n is large and p is not too small, in such a way that a Normal approximation is appropriate (this will usually be the case if $n \geq 50$ & $np \geq 10$), then X approx. $\sim N(np, np(1-p))$.

[A continuity correction is not usually required.]

So the test statistic is X, the number of successes, and the critical value is obtained from the Normal table. [Note: For the earlier tests, a sample is required; but here we are using a single value resulting from n trials.]

[Alternatively, $Y = \frac{X}{n}$ can be taken to be the test statistic, and use made of the fact that Y approx. $\sim N\left(p, \frac{p(1-p)}{n}\right)$: this approach is used for determining a confidence interval for the population proportion (see "Confidence Intervals").]

(8) Test for mean of a Poisson distribution

Option 1: $H_0: X \sim Po(\lambda)$

Test statistic: X, with the critical value obtained from the cumulative Poisson table (or calculated manually).

[Note: As for the Binomial proportion, we are using a single value - effectively from an infinite number of trials, with an infinitesimal probability of success.]

Option 2: $H_0: X \sim Po(\lambda)$ approx. $\sim N(\lambda, \lambda)$

Test statistic: X, with the critical value obtained from the Normal table. X should be >10. [A continuity correction is not usually required.]

(9) Test for variance of a Normal distribution

Option 1

$H_0: X \sim N(\mu, \sigma^2) \Rightarrow \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

Thus the test statistic is $X^2 = \frac{(n-1)S^2}{\sigma^2}$, with the critical value obtained from the χ^2 table.

Option 2

$H_0: X \sim N(\mu, \sigma^2) \Rightarrow \frac{S^2}{\sigma^2} \sim F_{n-1,\infty}$

Thus the test statistic is $\frac{S^2}{\sigma^2}$, with the critical value obtained from the F table.

(B) Paired samples

(10) Test for difference of two means from Normal distributions

Treat differences as a single sample, determining their mean and variance, and proceed as for (A)(1)-(4). [In practice, the paired samples are likely to be small, with the population variance unknown, so that a t-test will be needed.]

(11) Test for difference between two unknown distributions, where the paired samples are small

H_0: The two samples are from a common distribution

Apply either the Wilcoxon signed-rank test or the Sign test to the differences (as in (5) and (6)), with $M = 0$, according to whether the distribution can be assumed to be symmetrical or not.

(C) Two independent samples

(12) Test for given difference between means of Normal distributions, with known common variance

$H_0: \overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2})$

Test statistic: $z = \frac{(\overline{x}-\overline{y})-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

Note: Often $\mu_1 - \mu_2 = 0$

Variations (one or more of the following):

(i) Unknown common variance, with large samples: use s^2 (estimated from pooled data: $s^2 = \frac{(n_1-1){s_1}^2+(n_2-1){s_2}^2}{n_1+n_2-2}$; this is an unbiased estimator)

(ii) Unknown common variance, and small samples (from Normal distributions): apply t-test with $s^2 = \frac{(n_1-1){s_1}^2+(n_2-1){s_2}^2}{n_1+n_2-2}$

and $n_1 + n_2 - 2$ degrees of freedom

(iii) Unknown distributions, and large samples, so that approximate Normal distributions can be assumed, by the Central limit theorem

(iv) Different variances: replace $\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}$ with $\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}$

(${\sigma_1}^2$ and ${\sigma_2}^2$ can be estimated by ${s_1}^2$ and ${s_2}^2$, if the samples are large)

(13) Test for difference between two unknown distributions, where the (independent) samples are small

Apply the Mann-Whitney test (related to the Wilcoxon Rank Sum test [see below] - not to be confused with the Wilcoxon Signed Rank test, used earlier (especially as both methods involve summing ranks).

Procedure

(i) H_0: The two samples are from a common distribution

(ii) Suppose that the sample sizes are m & n, where $m \leq n$

(iii) Rank the items in both samples together, with a rank of 1 for the smallest value. For example, one sample (of size 5) may contain the ranks 2, 4, 7, 11, 13; whilst the other (of size 8) has the ranks 1, 3, 5, 6, 8, 9, 10, 12.

(iv) The test statistic is $U = T - \frac{1}{2}m(m+1)$,

where T is the sum of ranks for the smaller sample (or either if $m = n$) [$\frac{1}{2}m(m+1) = 1 + \cdots + m$ is the smallest possible value for T, so that $U \geq 0$]

[Note: The Wilcoxon Rank Sum test has as its test statistic the T above, so that the critical values are those of the Mann-Whitney test, with $\frac{1}{2}m(m+1)$ added.]

(v) Reject H_0 if U $<$ the lower tail critical value from the Mann-Whitney table [ie if U is suspiciously small if H_0 is assumed]

[Note: For larger m & n, $U \sim$ approx. $N(\frac{1}{2}mn, \frac{1}{12}mn(m+n+1))$; a continuity correction should be applied.]

(14) Test for equality of variance of two Normal distributions (given two independent samples)

Test statistic: $F = \frac{{s_1}^2}{{s_2}^2}$, where ${s_1}^2 > {s_2}^2$

Reject H_0 if F > upper tail critical value of F_{n_1-1,n_2-1}

Note: To test for a given ratio ${\sigma_1}^2/{\sigma_2}^2$ of variances, test statistic becomes $F = \frac{{s_1}^2/{\sigma_1}^2}{{s_2}^2/{\sigma_2}^2}$

(D) More than two independent samples

(15) Kruskal-Wallis test

Procedure:

(i) H_0: The samples are from a common distribution

(ii) Let the sample sizes be n_i, where $i = 1$ to k (and there are k samples), and let $N = \sum_{i=1}^{k} n_i$

(iii) Rank the items in all the samples together, with a rank of 1 for the smallest value. For example, one of the samples (of size 5) may contain the ranks 2, 4, 7, 11, 13

(iv) The test statistic is $H = [\frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{{T_i}^2}{n_i}] - 3(N+1)$,

where T_i is the sum of ranks for the ith sample

(v) Reject H_0 if H > upper tail critical value of ${\chi^2}_{k-1}$