CA200 – Quantitative Analysis for Business Decisions

File name: CA200_Section_05B_Statistical InferenceII

Table of Contents

5.6 Modifications: One and Two Samples; 1-Sided, 2-Sided Tests

Based on the properties of the sampling distributions introduced earlier, i.e. the Normal, Student 't' , χ^2 and F, hypothesis testing ideas *generalise.*

- For example, as before, if sample size **n < 25**, and we use the sample variance, and want to test a hypothesis about **the mean** of a sample, we use the $\mathbf{t_{n-1}}$ **distribution**
- Can also vary level of significance (risk) of test , (α) , where the usual values used are 0.05 (5%) or less.
- Can also apply tests : to proportions (use Normal)

  : to **differences** between means or proportions (a bit more work needed on the standard error term, but **Normal** or **t** as before)

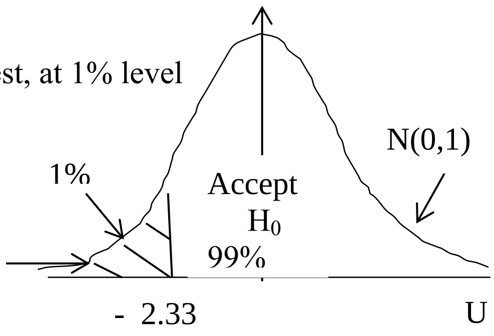
  : to a **single** variance or **ratio** of variances, (need other distributions i.e. χ^2 and **F** respectively).
- Can also look for differences in a given *direction***.** (Examples: (i) It may be acceptable for mean product weight to be higher than expected, for example, but not for it to be lower. (ii) A company may claim that more than 80% of personnel travel less than 30 km. to work. So, interested in departures which indicate that this is *not* the case, i.e. that the proportion travelling more than 30 km. is much more than 0.2 (or proportion travelling 30km or less is much less than 0.8) . In these cases, as usual, the *null hypothesis* is that statement, which means that everything is *operating as expected*. The *alternative hypothesis* is that statement which says it is *not*.

  So, for the second example, looking for a 1-sided test, at 1% level

  H_0: $p \geq 0.80$ (proportion at least 0.8 or 80%)

  H_1: $p < 0.8$ (proportion less than 0.8)

  Rejection region: 1% - to one side



Distribution: Proportions, so Normal.
Decision Rule:
Accept H_0, if U value generated on basis of sample value, (using the U-transform as usual), is greater than or equal to - 2.33.
Reject H_0 if the value of U generated by the sample is less than – 2.33.

5.7 H.T. Examples

Example 7

In a poll, 40% of a random sample of 1000 people indicate satisfaction with govt policy. Test at the *0.01 level of significance (1% risk level)* if this is **consistent** with the government claim of 45% support from the public.

The **claim** for the population proportion, P is that P = 0.45 . Sample gives sample proportion, p = 0.40, for a sample size, n = 1000

So $H_0 : P = 0.45$
$H_1 : P \neq 0.45$

- Proportions, so use **Normal Distribution**
- **Large** sample**,** effectively **infinite population,** so do *not* need **fpc**

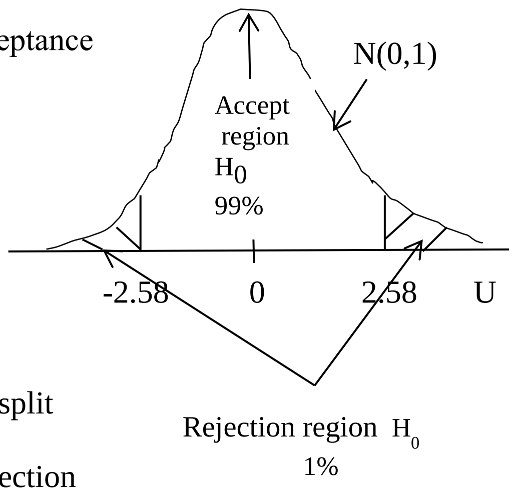
Standard Error (S.E.(proportion) is $\sqrt{pq/n}$ as before , so here have $\sqrt{(0.4)(0.6)/1000} = 0.015$

Test statistic for Normal in general is : $U = \dfrac{(\bar{x} - \mu)}{s/\sqrt{n}}$ so here $U = \dfrac{(p-P)}{\sqrt{pq/n}} = \dfrac{(0.4 - 0.45)}{0.015} = -3.33$

Critical values dividing up distribution into 99% and 1% are ± 2.58

Decision Rule:
Accept H_0 if Test Statistic value, based on sample, falls into Acceptance (or non-rejection) Region for H_0
Reject H_0 if this T.S. value falls into Rejection Region for H_0

Clearly, -3.33 is beyond -2.58, so Reject H_0 at 1% level of Significance. The govt. claim appears to be false.



Note: this is a 2-sided (or 2-tailed) test as the rejection region is split between the two tails of the distribution. This reflects the fact that the alternative hypothesis implies interest not just in one direction of departure from (or evidence against) H_0 but in either direction.

One-Tailed Tests

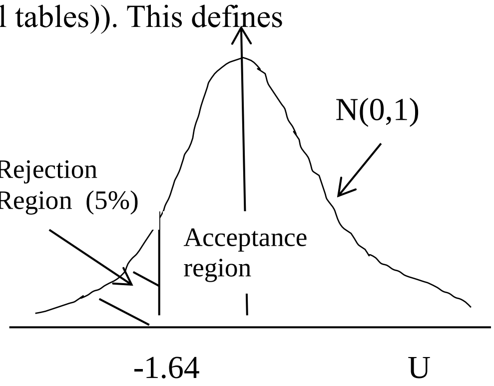
Suppose now that H_0 is of the form $\mathbf{H_0 : P \geq 0.45}$, then arbitrary *large* values of p (sample proportion) are *acceptable* under claim or statement that the null hypothesis is making, (i.e. do not invalidate it), but small values do provide evidence *against* H_0, i.e. lean towards the alternative hypothesis that $\mathbf{H_1 : P < 0.45}$

In this case the *direction of interest* is when a **smaller value than 45%** ($p < 0.45$) is returned from a sample, so, in this case, the **rejection region** for the test statistic lies in the *left hand tail* or *left-hand side* of distribution only.

<u>Example 8</u>

40% of a random sample of 1000 people in the country indicate satisfaction with govt. policy. Test at the $\alpha = 0.05$ *level of significance* if sample results are consistent with the claim that *at least 45%* support govt. policy?

Most of the calculation is the same, but now the distribution is divided up, s.t. 5% lies in the *left-hand tail*, (corresponding to a critical value of **-1.64** (see Normal tables)). This defines the 5% *one-sided rejection region* for H_0



The U-value obtained before was -3.33.
Clearly this falls well into the one-sided rejection region, so Reject H_0. The evidence is *against* 45% or more of the public supporting the govt. (at 5% risk level or level of significance $\alpha=0.05$).

<u>Example 9</u>

Based on past experience a company knows that, if the population S.D. of tyre life is **no more than** 3500 miles, then the *production process is working correctly.*
Given different shifts, the production manager is interested in determining if the Friday shift leads to **poorer quality,** (i.e. **S.D. larger**).

A random sample of 16 tyres, produced on Friday, gives the following results for life-testing

Tyre life (x) in thousands of miles)

22.42, 18.36, 21.46, 19.20, 23.40, 27.38, 23.46, 23.51, 20.62, 26.47, 19.75, 20.30, 17.84, 26.34, 28.27, 24.73

For these data, we want to calculate the **sample variance** as usual. A slightly easier way to perform these calculations requires a re-arrangement of the usual expression, to give:

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$$

<u>Solution</u>

The rearrangement of the expression for the estimated sample variance means that we can now tabulate the calculations, so that we are effectively summing, and squaring and summing, component values, as shown:

Sample Item	x_i	x_i^2
1	22.42	502.656
2	18.36	337.090
3	21.46	460.532
4	19.20	368.640
5	23.40	547.560
6	27.38	749.664
7	23.46	550.372
8	23.51	552.720
9	20.62	425.184
10	26.47	700.661
11	19.75	390.063
12	20.30	412.090
13	17.84	318.266
14	26.34	693.796
15	28.27	799.193
16	24.73	611.573
Sum	**363.51**	**8420.06**

Substituting in expression given, with n= 16 (sample size) gives:

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} = \frac{16(8420.06) - (363.51)^2}{16(15)} = 10.756$$

So standard deviation s = 3.2796 (thousands of miles)

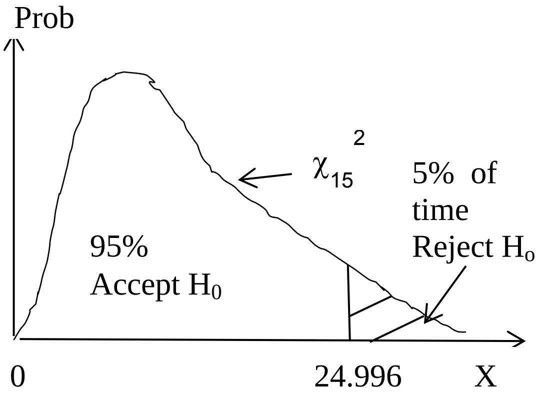
The **Hypotheses** are:

$H_0 : \sigma \leq 3500$ (or $H_0 : \sigma^2 \leq 12{,}250{,}000$)

$H_1 : \sigma > 3500$ (or $H_1 : \sigma^2 > 12{,}250{,}000$)

Distribution Choice. Testing a *single variance*, so **Chi-squared distribution** on 15 degrees of freedom (i.e. **n-1**), where the d.o.f. are based on the sample variance (sum of squared deviations from mean divided by n-1) as for t-distribution.

Critical value (from statistical tables = 24.996 to divide distribution as shown, because interested in one-sided alternative, i.e. higher values of S.D.)



So have

Decision Rule:

Accept H_0 if value obtained from sample ≤ value from χ^2_{15} that divides up distribution into 95% (Acceptance or non-rejection region), and 5% rejection region as shown. Equivalent to *level of significance* or *risk* = α = 0.05 (or 5%))

Reject H_0 if value obtained from sample is greater than 24.996, i.e. falls into rejection region.

Calculating Test Statistic

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$= \frac{(16-1)10755900}{(3500)^2} = 13.17$$

Clearly sample value is less than the critical value, so do not reject H_0. *No evidence* (at 5% level) *against* null hypothesis, i.e. no indication that Friday shift produces poorer quality tyres

Example 10.

Suppose that in needing to make a decision on variability of tyre life produced on Fridays generally, the production manager also wants to know if Friday afternoon production is worse than Friday morning. Of the 16 tyres sampled, 7 were taken from the morning shift, 9 from the afternoon, so the original data are grouped as:

Shift	Tyre Life (000s of miles)							
a.m.	22.42	18.36	21.46	19.20	23.40	27.38	23.46		
p.m.	23.51	20.62	26.47	19.75	20.30	17.84	26.34	28.27	24.73

With summarised calculations:

a.m.	p.m.
$\sum x_i = 155.68$	$\sum x_i = 207.83$
$\sum x_i^2 = 3516.514$	$\sum x_i^2 = 4903.545$
$n_{a.} = 7$	$n_p = 9$
d.o.f. = 7-1 =6	d.o.f. = 9-1= 8

Substituting in the expression for variance, as before, then we obtain the following for the two variances (i.e. for a.m. and p.m.)

$$s_{am}{}^{2} = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} = \frac{7(3516.514)-(155.68)^2}{7(6)} = 9.032$$

$$s_{pm}{}^{2} = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} = \frac{9\,(4903.545)-(207.83)^2}{9(8)} = 13.036$$

Note that this looks, superficially, as though there is increased variability, but may not be significant, so must test.

Hypotheses:

$H_0 : \sigma_{am}^2 \le \sigma_{pm}^2$

$H_1 : \sigma_{am}^2 > \sigma_{pm}^2$

Distribution:

Ratio of two variances, so **F-distribution** on d.o.f. given by s_{am}^2 and s_{pm}^2

Decision Rule :

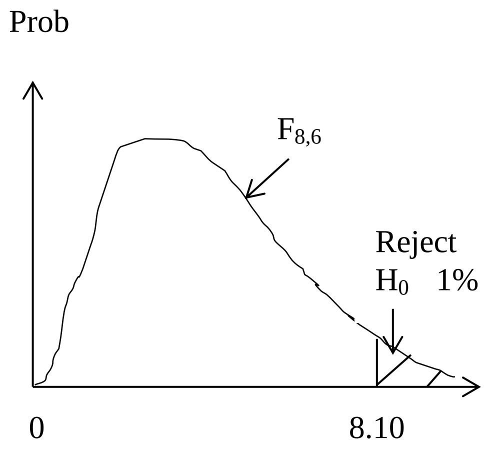
Accept H_0 if ratio of variances ≤ distribution value at 1% level (say).

[Choose own *level of significance*, α, i.e. *risk level* if not given it].

Reject H_0 if ratio of variances > distribution critical value at 1% level (say).

So, calculation gives us:

$$F_{(n_{pm}-1),(n_{am}-1)} = \frac{s_{pm}{}^{2}}{s_{am}{}^{2}} = \frac{13.036}{9.032} = 1.443$$



Where the largest variance is usually set as the numerator for convenience of reading F-tables: (a ratio so does not matter)

Obviously can *not* reject H_0., as 1.443 much less than 8.10, i.e. Falls well within acceptance region for null hypothesis.
No evidence (at *1%* level of significance), that production is worse: (more variable tyre life) from the p.m. shift.

Differences between Means

Typically, interested in *ratios* of variances (as squared) and *differences* in means or proportions. For proportions, we always use the **Normal distribution** and adjust the standard error to reflect that two proportions are involved. For differences in means, we can use the Normal if the samples are large and variances known, but a more conservative test is based on the t-distribution, and allows for small samples as well, - as follows:

Suppose that values x_1 x_2 … x_m are a random sample, with mean $\bar{x}$ and standard deviation s_1 drawn from a distribution with mean μ_1 and variance σ_1^2 unknown but estimated by sample variance. Similarly, y_1 y_2 … y_n is a random sample , mean $\bar{y}$ and standard deviation s_2 drawn from distribution with mean μ_2 and also uses its sample variance.

A *typical test of interest* is to see if both samples are effectively drawn from the same parent population, i.e. have *equal means*.

Hypotheses:

H_0: $\mu_1 = \mu_2$ or H_0: $\mu_1 - \mu_2 = 0$

H_1: $\mu_1 \neq \mu_2$ H_1: $\mu_1 - \mu_2 \neq 0$

The *pooled estimate* of the parent population variance weights both independent variances by their sample size

$$s^2 = \frac{(m-1)s_m^2 + (n-1)s_n^2}{(m-1)+(n-1)}$$

Where this is the (conservative) estimate of the variance of the *difference of two means*

This gives a Test Statistic, which has a $\mathbf{t_{m+n-2}}$ **distribution**, where the *degrees of freedom* from the test statistic come from adding the degrees of freedom of the *two variances* involved, so Test Statistic

$$T_{m+n-2} = \frac{[(\bar{x}-\bar{y}) - (\mu_1 - \mu_2)]}{s\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)}}$$

The next example illustrates how this is applied.

Example 11

A random sample size m = 25 has mean $\bar{x} = 2.5$ and standard deviation $s_m = 2$.
A second random sample, size n = 41 has mean $\bar{y} = 2.8$ and standard deviation $s_n = 1$.
Test at the 0.05 level of significance if means of the parent populations are the same.

Hypotheses: $H_0 : \mu_1 = \mu_2$ or H_0: $\mu_1 - \mu_2 = 0$

$H_1 : \mu_1 \neq \mu_2$ H_1: $\mu_1 - \mu_2 \neq 0$

Difference in sample means is the ‘sample statistic’ in this case i.e. $\bar{x} - \bar{y}$

The pooled estimate of variances is : $s^2 = \frac{(m-1)s_m^2 + (n-1)s_n^2}{(m-1)+(n-1)} = 2.125$

so Test Statistic is $T_{m+n-2} = \frac{[(\bar{x}-\bar{y})-(\mu_1-\mu_2)]}{s\sqrt{\left(1/m + 1/n\right)}} = \frac{(\bar{x}-\bar{y})-0}{s\sqrt{\left(1/m + 1/n\right)}} = \frac{-0.3}{1.457\sqrt{\left(1/25 + 1/41\right)}} = -0.811$

We set up the decision rule as usual and note here that the hypotheses are specified such that *either direction of departure* from the claim is possible, i.e. if the means are not equal then the mean of population one, μ_1, (estimated by $\bar{x}$) may be greater than the mean of population 2, μ_2 , (estimated by $\bar{y}$), or *vice versa* so the t- distribution is divided up s.t. 95% contained between the critical values and 2½% in each tail.

From the tables, the 0.05 critical value for the t-distribution on 64 d.o.f. is approximately 2: (check as exercise – approaching limit of tables).

Clearly, as -0.811 falls between ±2, i.e. is within the acceptance region, we can *not* reject the null hypothesis at the 5% level of significance, so we can treat the means as equal.