

Displays for Statistics 5401/8401

Lecture 9

September 26, 2005

Christopher Bingham, Instructor

612-625-1024, kb@umn.edu

372 Ford Hall

Class Web Page

<http://www.stat.umn.edu/~kb/classes/5401>

© 2005 by Christopher Bingham

I did a small simulation experiment to accomplish two purposes:

- Examine the *power* of the Q-Q $\sqrt{\chi^2}$ correlation test for normality
- Check *robustness* to non-normality of the small sample Hotelling's T^2 distribution. I used $M = 5000$ trials, with $n = 50$ and $p = 4$.

In each table there are four lines

Intended α The significance level used in the normality test or T^2 test

Power = power of the normality test with significance level α .

α with $F_{4,46} = \text{actual } \alpha \approx$

$P(\text{reject } H_0: \mu = 0) \text{ using } T^2 \text{ with small } n \text{ critical value } ((f_e p)/(f_e - p + 1)) F_{p, f_e - p + 1}(\alpha).$

α with $\chi_4^2 = \text{actual } \alpha \approx P(\text{reject } H_0: \mu = 0) \text{ using } T^2 \text{ with large } n \text{ critical value } \chi_4^2(\alpha)$

1. x_1, x_2, x_3, x_4 independent Student's t_5

Intended α	.01	.02	.05	.10
Power	.2282	.2902	.3982	.5042
α with $F_{4, 46}$.0068	.0162	.0446	.0996
α with χ_4^2	.0206	.0356	.0778	.1412

Distribution of t_5 has symmetric bell-shaped density but with thicker tails than normal density.

Power is moderate, small sample actual is reasonably close to intended but large sample actual α is not.

2. Independent Student's $t_3, \mu_0 = 0$.

Intended α	.01	.02	.05	.10
Power	.5560	.6328	.7490	.8250
Actual α	.0064	.0154	.0422	.0958
α with χ_4^2	.0174	.0336	.0772	.1388

t_3 is less normal than t_5 and power of correlation test is larger; actual small sample α is a little worse than for t_5 .

3. Independent $\chi_{10}^2, \mu_0 = \text{rep}(10, 4)$.

Intended α	.01	.02	.05	.10
Power	.129	.1852	.2886	.3884
Actual α	.016	.027	.0522	.1022
α with χ_4^2	.0306	.0438	.083	.1418

χ_{10}^2 is quite skewed with mean 10.

Power is less than for t_5 , actual α is not bad, a little larger than intended.

4. Independent $\chi_4^2, \mu_0 = \text{rep}(4, 4)$.

Intended α	.01	.02	.05	.10
Power	.3306	.4304	.5952	.7142
Actual α	.0206	.0332	.067	.1214
α with χ_4^2	.0396	.0574	.1016	.160

χ_4^2 is more skewed with mean 4.

Power is larger than for χ_{10}^2 and t_5 , actual α is double the intended α for $\alpha = .10$ and somewhat too large for smaller α .

Test of a vector parameter

Problem: Test $H_0: \boldsymbol{\theta} = \boldsymbol{\theta}_0$, where $\boldsymbol{\theta}$ is a vector of q parameters estimated by $\hat{\boldsymbol{\theta}}$, with $V[\hat{\boldsymbol{\theta}}]$ consistently estimated by $\hat{V}[\hat{\boldsymbol{\theta}}]$.

If $\mathbf{C} = \hat{V}[\hat{\boldsymbol{\theta}}]^{1/2}$ is a square root of $\hat{V}[\hat{\boldsymbol{\theta}}]$ (i.e., $\mathbf{C}'\mathbf{C} = \hat{V}[\hat{\boldsymbol{\theta}}]$) you might hope to base a test on the multistandardized statistic

$$\mathbf{Z} = (\mathbf{C}^{-1})'(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = (\{\hat{V}[\hat{\boldsymbol{\theta}}]\}^{-1/2})'(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$$

When $\hat{\boldsymbol{\theta}} \approx N_q(\boldsymbol{\theta}_0, V[\hat{\boldsymbol{\theta}}])$, $\mathbf{Z} \approx N_q(\mathbf{0}, \mathbf{I}_q)$.

Because $\hat{V}[\hat{\boldsymbol{\theta}}]^{1/2}$ is *not* unique, this is problematical. However, the statistic

$$T^2 = \|\mathbf{Z}\|^2 = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)'(\hat{V}[\hat{\boldsymbol{\theta}}])^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$$

is unique.

This forms the basis for many tests in both univariate and multivariate analysis.

- 1-way ANOVA F-test ($\boldsymbol{\theta} = [\mu_1, \dots, \mu_g]'$, $\hat{\boldsymbol{\theta}} = [\bar{x}_{.1}, \bar{x}_{.2}, \dots, \bar{x}_{.g}]'$, $\hat{V}[\hat{\boldsymbol{\theta}}] = \text{diag}\{s^2/n_i\}_{1 \leq i \leq g}$)
- $\chi^2 = \sum_i (O_i - E_i)^2 / E_i$ goodness of fit test.

Comment: There are several statistics called Hotelling's T^2

- one-sample Hotelling's T^2 ($H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$)
- two-sample Hotelling's T^2 ($H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$)
- paired Hotelling's T^2 ($H_0: \boldsymbol{\mu}_d = \mathbf{0}$), etc.

They may differ in

- The dimension p = number of means being tested.
- The error degrees of freedom f_e .
 f_e is usually the same as for the analogous Student's t degrees of freedom

Expressed in terms of p and f_e , *with normal data*, their small sample distributions are all the same:

$$T^2 = (pf_e / (f_e - p + 1)) F_{p, f_e - p + 1}$$

To use this you must know 2 numbers:

- p = dimension. This is usually obvious.
- f_e = error d.f. You can usually use f_e used in univariate test.

There are special formulas for particular cases, but I recommend you *don't* use them.

Examples:

For a one-sample T^2 , when you plug in $f_e = n - 1$ into

$$T^2 = \{pf_e / (f_e - p + 1)\} F_{p, f_e - p + 1}$$

you get the "special" formula

$$T^2 = \{p(n-1) / (n-p)\} F_{p, n-p}$$

In the two-sample case, $f_e = n_1 + n_2 - 2$.

When you plug this into the general form you get the "special" formula

$$T^2 = \{p(n_1 + n_2 - 2) / (n_1 + n_2 - p - 1)\} F_{p, n_1 + n_2 - p - 1}$$

Hotelling's one-sample

$$T^2 = (\bar{\mathbf{x}} - \boldsymbol{\mu}_0)' \hat{\mathbf{V}}(\bar{\mathbf{x}})^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$$

is one way to test $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$.

Another way is to use **Bonferroniized Student's t tests**.

$H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$ is equivalent to the p univariate null hypotheses $H_{0j}: \mu_j = \mu_{0j}$. You can test each of them with a t-statistic

$$t_1 = (\bar{x}_1 - \mu_{01}) / \sqrt{(s_{11} / n)},$$

$$t_2 = (\bar{x}_2 - \mu_{02}) / \sqrt{(s_{22} / n)}$$

...

$$t_p = (\bar{x}_p - \mu_{02}) / \sqrt{(s_{pp} / n)}$$

When H_{0j} is true t_j is Student's t on $f_e = n - 1$ d.f.

You can reject $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$ if any t_j is significantly large.

You have to be careful about how you determine which t_j are significant. If you just compare t_j with $t_{f_e}(\alpha/2)$ (two-tail tests), you will reject H_0 with probability substantially greater than α .

The solution is to *Bonferronize* the p t -tests. You do this by one of two ways:

- Use critical value $t_{f_e}((\alpha/p)/2)$
- Multiply each P-value by p

The resulting test of H_0 has true significance level $\leq \alpha$.

This is just as truly a multivariate test as is T^2 .

Because (actual α) \leq (intended α), this is a conservative test. In many cases, (actual α) \approx (intended α).

MacAnova example based on Fisher *Iris setosa* data

```
Cmd> y <- read("", "t11_05", quiet:T) # read JWData5.txt
Read from file "TP1:Stat5401:Data:JWData5.txt"
```

quiet:T prevents descriptive comments.

Test $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0 = [5.0, 3.4, 1.4, 0.2]'$:

```
Cmd> setosa <- y[y[,1]==1,-1] # extract setosa flower variables
Cmd> stats <- tabs(setosa, mean:T, covar:T) # compute
Cmd> compnames(stats) # names of components
(1) "mean"           1 by p Row vector
(2) "covar"         p by p matrix
Cmd> n <- nrows(setosa) # n = 50
Cmd> p <- ncols(setosa) # 4
Cmd> xbar <- stats$mean # column vector
Cmd> vhat <- stats$covar/n #  $\hat{V}[\bar{\mathbf{x}}] = \mathbf{S}/n$ 
Cmd> mu_0 <- vector(5.0, 3.4, 1.4, 0.2) # hypothesized value
Cmd> stderrs <- sqrt(diag(vhat)) # sqrt( $s[i,i]/n$ ),  $i=1, \dots, p$ 
Cmd> tstats <- (xbar - mu_0)/stderrs; tstats # white box
(1) 0.12036 0.52231 2.5245 3.0865
Cmd> tval(setosa - mu_0') # black box t-statistics
(1) 0.12036 0.52231 2.5245 3.0865
Cmd> fe <- n-1 # error d.f. 49
Cmd> pvalues <- p*twotailt(tstats, fe); pvalues #Bonferronized
(1) 3.6188 2.4152 0.059508 0.013314
Cmd> alpha <- .05; invstu((alpha/p)/2, fe, upper:T) #Bonf critval
(1) 2.5933 Bonferronized critical value
```

Conclusion: Since $t_4 = 3.0865 > 2.5933$ (or $pvalues[4] < .05$), you can reject the overall $H_0: \mu = \mu_0$ and the 5% level.

But you can also reject $H_{04}: \mu_4 = \mu_{04}$. Thus you have learned something about how the overall H_0 is untrue.

Testing using T^2

```
Cmd> tsq <- hotellval(setosa - mu_0'); tsq
(1,1)      13.616      Hotelling's T^2

Cmd> invchi(.05,p,upper:T) # upper tail probability point
(1)        9.4877      5% critical value for large n

Cmd> cumchi(tsq,p,upper:T) # upper tail probability
(1,1)      0.0086268      P-value for large n

Cmd> (fe*p/(fe - p + 1))*invF(.05, p, fe - p + 1, upper:T)
(1)        10.968      5% critical value for small n

Cmd> cumF(((fe-p+1)/(fe*p))*tsq,p,fe-p+1, upper:T)
(1,1)      0.02133      P-value for small n
```

Conclusion:

Since $T^2 = 13.616 > T_{.05}^2 = 10.968$ (or because $P = .02133 < .05$), you can **reject** $H_0: \mu = \mu_0$.

However, T^2 gives no information about how this overall H_0 is violated. This is a disadvantage of T^2 as compared to Bonferroni t-statistics.

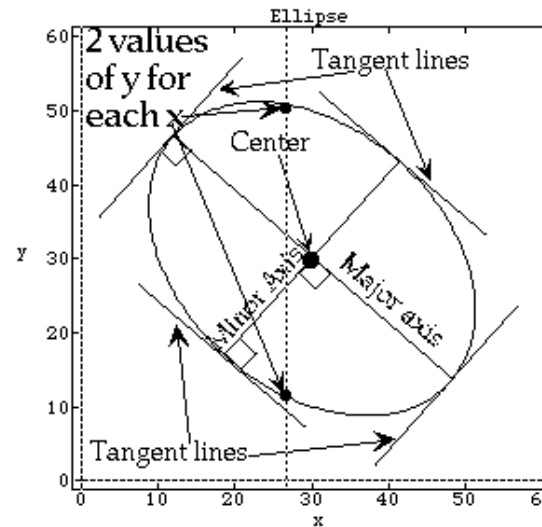
At this point it is sometimes suggested you do a "post hoc" analysis using t-statistics. But if you do that, you might as well start out with t-statistics.

Notes on MacAnova Computation

- `solve(a,b)` or `a \% \% b` computes $\mathbf{A}^{-1}\mathbf{b}$ and can be more accurate than `solve(a) \%* \% b`.
- $\text{cumchi}(x, df) = P(\chi_{df}^2 \leq x)$
 $\text{cumchi}(x, df, upper:T) = P(\chi_{df}^2 \geq x)$
 $P(\chi_{df}^2 \leq \text{invchi}(p, df)) = p$
 $P(\chi_{df}^2 \geq \text{invchi}(p, df, upper:T)) = p$
- `cumF(x, df1, df2 [, upper:T])` and `invF(p, df1, df2 [, upper:T])` compute lower tail [upper tail] cumulative probabilities and critical values for the F distribution.
- I prefer first to compute \hat{V} to use in $T^2 = (\bar{\mathbf{x}} - \boldsymbol{\mu}_0)' \hat{V} [\bar{\mathbf{x}}]^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$, rather than use a formula like $T^2 = n(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)' \mathbf{S}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$ n which $\hat{V}[\bar{\mathbf{x}}] = \mathbf{S}/n$ is sort of "hidden."

Ellipses and Ellipsoids

Here is an ellipse, a curve in the $p = 2$ dimensional x-y plane:



Major and minor axes are the longest and shortest "diameters" and are perpendicular to (\perp) tangent lines

Other diameters are not \perp tangent lines. All points (x,y) on an ellipse with center (x_0, y_0) satisfy an equation like

$$a(x-x_0)^2 + 2b(x-x_0)(y-y_0) + c(y-y_0)^2 = K^2$$

where a, b, c and K are constants with

- $a > 0, c > 0$ and $K > 0$
- $ac - b^2 > 0$

To get an expression for y in terms of x , you solve a quadratic equation to get:

$$y = f(x) \equiv y_0 - b(x-x_0)/c \pm \sqrt{\{K^2/c - (ac - b^2)((x-x_0)/c)^2\}}$$

You can't get a real square root of a negative number, so y is defined only for x that satisfy

$$K^2/c - (ac - b^2)((x-x_0)/c)^2 \geq 0,$$

that is, only for x that satisfy

$$|x-x_0| \leq K\sqrt{(c/D)}, \text{ with } D = ac - b^2 > 0$$



Define q^{11} , q^{12} , q^{21} , and q^{22} by

$$q^{11} \equiv a \quad q^{12} = q^{21} \equiv b \quad q^{22} \equiv c$$

In terms of q^{ij} , the ellipse equation is

$$q^{11}(x-x_0)^2 + 2q^{12}(x-x_0)(y-y_0) + q^{22}(y-y_0)^2 = K^2$$

with $q^{11} > 0$, $q^{22} > 0$, $D = q^{11}q^{22} - (q^{12})^2 > 0$.

Role of the constants

- The center of the ellipse is at (x_0, y_0) .
- For given q^{11} , q^{12} and q^{22} , the size of the ellipse is determined by K . A larger K produces a larger ellipse
- The shape of the ellipse is determined by the ratios q^{11}/K^2 , q^{12}/K^2 and q^{22}/K^2 .

Define the matrix

$$\mathbf{Q}^{-1} \equiv \begin{bmatrix} q^{11} & q^{12} \\ q^{12} & q^{22} \end{bmatrix}$$

Then

- the conditions $q^{11} > 0$, $q^{22} > 0$, $D = q^{11}q^{22} - (q^{12})^2 > 0$ are completely equivalent to \mathbf{Q}^{-1} being positive definite.

- $D = \det(\mathbf{Q}^{-1})$