Section 3: The Multivariate Normal

20 Sept. 01

Properties of the normal.

Definition of the multivariate normal, including singular case

$Z_1, \dots, Z_\ell \quad : \quad \text{I}\, N(0,1)$

$$\overset{\ell\times 1}{\underset{\sim}{Z}} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_\ell \end{bmatrix}$$

$$E\,\underset{\sim}{Z} = 0, \qquad \text{var}\,\underset{\sim}{Z} = \underset{\sim}{\Sigma}_{ZZ} = \underset{\sim}{I}$$

$$\overset{m\times 1}{\underset{\sim}{Y}} = \overset{m\times 1}{\underset{\sim}{\alpha}} + \overset{m\times \ell}{\underset{\sim}{\beta}}\,\overset{\ell\times 1}{\underset{\sim}{Z}} \rightleftharpoons N_m(\underset{\sim}{\mu}_Y, \underset{\sim}{\Sigma}_{YY})$$

$$\underset{\sim}{\mu}_Y = \underset{\sim}{\alpha}, \qquad \underset{\sim}{\Sigma}_{YY} = \underset{\sim}{\beta}\,\underset{\sim}{\beta}'$$

Theorem. $\overset{n\times 1}{\underset{\sim}{X}} = \overset{n\times 1}{\underset{\sim}{A}} + \overset{n\times m}{\underset{\sim}{B}}\,\overset{m\times 1}{\underset{\sim}{Y}}$ is $N_n(\underset{\sim}{A} + \underset{\sim}{B}\,\underset{\sim}{\mu}_Y, \underset{\sim}{B}\,\underset{\sim}{\Sigma}_{YY}\,\underset{\sim}{B}')$

Proof. Linear combination of Z's.

Dropping subscript Y 20 Sept 01

Theorem. The m.g.f.

$$E\{e^{\underset{\sim}{\theta}'\underset{\sim}{Y}}\} = e^{\underset{\sim}{\theta}'\underset{\sim}{\mu} + \frac{1}{2}\underset{\sim}{\theta}'\underset{\sim}{\Sigma}\underset{\sim}{\theta}}$$

Proof.

$$\underset{\sim}{\theta}'\underset{\sim}{Y} = \underset{\sim}{\theta}'\underset{\sim}{\alpha} + \underset{\sim}{\theta}'\underset{\sim}{\beta}\underset{\sim}{Z} = \underset{\sim}{\theta}'\alpha + \underset{\sim}{\delta}'\underset{\approx}{Z} \qquad \underset{\sim}{\delta} = \underset{\sim}{\beta}'\underset{\sim}{\theta}$$

$$= \gamma + \sum_i \delta_i Z_i$$

$$E\, e^{\underset{\sim}{\theta}'\underset{\sim}{Y}} = e^{\gamma} e^{\frac{1}{2}\sum\delta_i^2} = e^{\gamma} e^{\frac{1}{2}\underset{\sim}{\delta}'\underset{\sim}{\delta}}$$

$$= \exp\{\underset{\sim}{\theta}'\underset{\sim}{\alpha} + \frac{1}{2}\underset{\sim}{\theta}'\underset{\sim}{\beta}\underset{\sim}{\beta}'\underset{\sim}{\theta}]$$

Shows well defined wrt choices of β
– need $\underset{\sim}{\beta}\underset{\sim}{\beta}'$ unique

Corollary. $\underset{\sim}{\Sigma} = \begin{bmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{\Sigma}_{22} \end{bmatrix} \Rightarrow \underset{\sim}{Y}_1 \perp\!\!\!\perp \underset{\sim}{Y}_2$

$$\underset{\sim}{\theta}'\underset{\sim}{\Sigma}\underset{\sim}{\theta} = \underset{\sim}{\theta}_1'\underset{\sim}{\Sigma}_{11}\underset{\sim}{\theta}_1 + \underset{\sim}{\theta}_2'\underset{\sim}{\Sigma}_{22}\underset{\sim}{\theta}_2$$

i.e. mgf factors

20 Sept 04

Conditionals of the normal

$$\overset{r+s}{\begin{bmatrix} \underset{\sim}{Y}_1 \\ \underset{\sim}{Y}_2 \end{bmatrix}} \quad \text{mean} \quad \begin{bmatrix} \underset{\sim}{0} \\ \underset{\sim}{0} \end{bmatrix}$$

$$\text{covariance} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

(normal) equation $\overset{r\times s}{\Theta}\,\overset{s\times s}{\Sigma_{22}} = \overset{r\times s}{\Sigma_{12}}$

consider some solution eg $\Theta = \Sigma_{12}\Sigma_{22}^{-}$

Set $\overset{r\times 1}{\underset{\sim}{\epsilon}} = \overset{r\times 1}{\underset{\sim}{Y}_1} - \underset{\sim}{\Theta}\,\underset{\sim}{Y}_2$

ϵ is normal (linear combination of normals)

$E\epsilon = 0$

$\text{cov}\{\epsilon, Y_2\} = \Sigma_{12} - \Theta\Sigma_{22} = 0$

$\epsilon \perp\!\!\!\perp Y_2$

$$\text{var}\,\epsilon_1 = \Sigma_{11} - \Theta\Sigma_{21} - \Sigma_{12}\Theta' + \overbrace{\Theta\Sigma_{22}}^{\Sigma_{12}}\Theta'$$

$Y_1 = \Theta Y_2 + \epsilon$

$$= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21}$$

$\underset{\sim}{Y}_1 \mid \underset{\sim}{Y}_2 \ : \ N_r(\Sigma_{12}\Sigma_{22}^{-}Y_2,\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21})$

means not 0

$$N_r(\mu_1 + \Sigma_{12}\Sigma_{22}^{-}(Y_2 - \mu_2), \qquad)$$

Regression Case

$Y = X\beta + \epsilon$ $\qquad \epsilon$: normal

$\hat{Y} = X\hat{\beta} = HY$

$\hat{\epsilon} = Y - \hat{Y} = (I-H)Y$

$\hat{Y} : N_{\sim}(X\beta, H\sigma^2)$ $\qquad$ rank(H) = r so need a g-inv

$\hat{\epsilon} \perp\!\!\!\perp \hat{Y}$ $\qquad$ saw $\perp$ earlier

$\hat{\epsilon} : N_{\sim}(0, (I-H)\sigma^2)$

$P'\hat{\beta} = P'(X'X)^{-}X'Y$

$P'\hat{\beta} : N(P\beta, P'\underset{-1}{(X'X)^{-}}P\sigma^2)$

(*) $\hat{\epsilon} \perp\!\!\!\perp \underset{\sim}{P}\underset{\sim}{\hat{\beta}}$ $\qquad$ saw $\perp$ earlier

What about $\hat{\sigma}^2$?

$\hat{\sigma}^2 = \underset{\sim}{\epsilon}'(\underset{\sim}{I} - \underset{\sim}{H})\underset{\sim}{\epsilon} / (n-r)$

quadratic form in the ϵ's

$\underset{\sim}{\epsilon}'\underset{\sim}{A}\underset{\sim}{\epsilon}$ $\qquad \underset{\sim}{A}$: symmetric, idempotent

$r(I-H) = tr(I-H) = n-r$

20 Sept 01

Some related distributions

$Z_1, \dots, Z_f$: IN$(0,1)$

chi-squared

$$\chi_f^2 = Z_1^2 + \dots + Z_f^2$$

Students t with f degrees of freedom

$$t = Z / \sqrt{\chi_f^2 / f}$$

Z and χ_f^2 independent

F with degrees of freedom f_1 and f_2

$$F = \frac{\chi_{f_1}^2 / f_1}{\chi_{f_2}^2 / f_2}$$

⑤ 20 Sept. 01

Theorem

$$\underset{\sim}{Y}_2 : N_s(\underset{\sim}{0}, \overset{s\times s}{\underset{\sim}{\Sigma}})$$

$$\overset{r\times 1}{\underset{\sim}{Y}_1} = \overset{s\times r}{\underset{\sim}{B}}\,\overset{r\times 1}{\underset{\sim}{Y}_2} : N_r(\underset{\sim}{0}, \underset{\sim}{B}'\underset{\sim}{\Sigma}\underset{\sim}{B})$$

Suppose $\underset{\sim}{\Sigma} = \sigma^2 \underset{\sim}{I}$ and $\underset{\sim}{B}'\underset{\sim}{B} = \underset{\sim}{I}$, then

$$\underset{\sim}{Y}_1 : N_r(\underset{\sim}{0}, \sigma^2 \underset{\sim}{I})$$

In particular orthogonal transformation $IN(0,\sigma^2)$ gives $IN(0,\sigma^2)$

---

this $\underset{\sim}{\epsilon}'\underset{\sim}{A}\underset{\sim}{\epsilon}$ $\qquad \epsilon : N_n(0, \sigma^2 I)$

Lemma: If A is symmetric, idempotent, then $\lambda_j = 1$ or 0

svd $\quad A = O'\Lambda O$

$$O'\Lambda O O'\Lambda O = O'\Lambda O$$
$$O'\Lambda^2 O = O'\Lambda O$$
$$\Lambda^2 = \Lambda$$
$$\lambda_j = {1 \atop 0}$$

Theorem

$$\underset{\sim}{\epsilon}'\underset{\sim}{A}\underset{\sim}{\epsilon} = \underset{\sim}{\epsilon}'\underset{\sim}{O}'\underset{\sim}{\Lambda}\underset{\sim}{O}\underset{\sim}{\epsilon} = \underset{\sim}{\eta}'\underset{\sim}{\Lambda}\underset{\sim}{\eta}$$

$$\underset{\sim}{\eta} : N(\underset{\sim}{0}, \sigma^2 \underset{\sim}{I})$$

$$\underset{\sim}{\epsilon}'\underset{\sim}{A}\underset{\sim}{\epsilon} = \sigma^2 \chi^2_f \qquad f = \text{rank}(\underset{\sim}{A}) = \text{tr}(\underset{\sim}{A})$$

20 Sept 01

Why [illegible] , can $p'\hat{\beta}$ [illegible] ? [illegible]

$$\hat{\sigma}^2 \sim \sigma^2 \chi^2_{(n-r)} / (n-r)$$

and $\perp\!\!\!\perp$ $\underset{\sim}{p}'\hat{\underset{\sim}{\beta}}$, $\underset{\sim}{X}\hat{\underset{\sim}{\beta}}$ $\qquad \hat{\underset{\sim}{\varepsilon}} \perp\!\!\!\perp \underset{\sim}{p}'\hat{\underset{\sim}{\beta}}$ (X)

Confidence intervals for $\underset{\sim}{p}'\hat{\beta}$

$$t = \frac{p'(\hat{\beta}-\beta)}{\hat{\sigma}\sqrt{p'(X'X)^{-}p}}$$ students t $\quad$ n-r degrees of freedom

$$= \frac{p'(\hat{\beta}-\beta)/\sqrt{p'(X'X)^{-}p}}{\hat{\sigma}}$$

$$= \frac{\sigma z}{\sigma\sqrt{(z_1^2+\dots+z_{n-r}^2)/(n-r)}}$$

$$t^2 = F_{1,n-r} \qquad \boxed{\text{p-value, test } \beta = \beta_0}$$

Suppose $\underset{\sim}{p}$ is an $m \times k$ matrix
and wish ~~examine hypothesis~~ $\underset{\sim}{p}'\underset{\sim}{\beta} = \underset{\sim}{\theta}$ $\qquad$ several relations

~~testable~~ hypothesis $\equiv$ estimable function

20 Sept 01

$$\underset{\sim}{p}'\hat{\underset{\sim}{\beta}} : N_k(\underset{\sim}{p}'\underset{\sim}{\beta}, \sigma^2 \underset{\sim}{p}'(\underset{\sim}{X}'\underset{\sim}{X})^- \underset{\sim}{p})$$

$$\underset{\sim}{p}'\hat{\underset{\sim}{\beta}} \perp\!\!\!\perp \|Y - X\hat{\beta}\|^2$$

$$(\underset{\sim}{p}'\hat{\underset{\sim}{\beta}} - \underset{\sim}{\theta})'[\underset{\sim}{p}'(\underset{\sim}{X}'\underset{\sim}{X})^- \underset{\sim}{p}]^-(\underset{\sim}{p}'\hat{\underset{\sim}{\beta}} - \underset{\sim}{\theta}) : \sigma^2 \chi^2_q$$

Source of this result? —use svd form

$$q = \text{rank}(\underset{\sim}{p})$$

$$F_{q,n-r} \sim \frac{SSH/q}{SSE/(n-r)}$$

ratio indep't χ^2

Numerator and denominator independent from ⊛ way above.

20 Sept 01

Fisher - Cochran Theorem (Rao, p. 185) $\underset{\sim}{Y} : N(0, \sigma^2 \underset{\sim}{I})$

Let $Q_1, \dots, Q_k$ be k quadratic forms ($Q_i = \underset{\sim}{Y}' \underset{\sim}{A}_i \underset{\sim}{Y}$) with ranks $f_1, \dots, f_k$ such that

$$\underset{\sim}{Y}' \underset{\sim}{Y} = Q_1 + \dots + Q_k .$$

Then a n.s. condition that $Q_i \sim \chi^2_{f_i}$ and are independent is $n = f_1 + \dots + f_k$.

Proved by decomposing $\underset{\sim}{Y}$ into independent variates going along with the Q_i.

Don't need $\underset{\sim}{A}_i \geq 0$

$$Q_i = \sum_{\alpha \in S_i} \pm (\underset{\sim}{B}_\alpha \underset{\sim}{Y})^2$$

$$\sum_{\alpha \in S_i} 1 = n_i$$

$$\sum_i Q_i = \underset{\sim}{Y}' \underset{\sim}{B}' \underset{\sim}{\Delta} \underset{\sim}{B} \underset{\sim}{Y} , \quad \underset{\sim}{\Delta} = \text{diag}\{\pm 1\}$$

25 Sept 01

<u>The birth data</u>. Now we know about the normal!

ANOVA of $\sqrt{4*\text{count}+1}$

Source	SS	DF	MS	F	p-value
Between days of week	106.75	6	17.79	14.81	5.55 E-16
Residual	857.66	714	1.20		
Total	964.41	720			

Model $Y_{jk} = \mu_j + \epsilon_{jk}$ $\epsilon_{jk} : I\,N(0, \sigma^2)$

$H_0 : \mu_1 = \mu_2 = \ldots = \mu_7$ $J = 7, K = 103$ (weeks)

$\mu_1 - \mu_2, \mu_1 - \mu_3, \ldots, \mu_1 - \mu_7 = 0$

P-value: probability $F \geq 14.81$ under H_0

$= 1 - pf(14.81, 6, 714) = 5.55 E{-16}$

Either an event of very (very) small probability has occurred or H_0 is unreasonable.

¿ Source of small P-value?

25 Sept 01

The problem of multiplicity

Consider the model

$$Y_{jk} = \mu_j + \epsilon_{jk} \qquad \epsilon_{jk} : IN(0, \sigma^2)$$

and suppose one wonders whether the

$$\mu_j = \mu_0 \ , \ \mu_0 \text{ given} \quad j = 1, .., J$$

Suppose σ is known.

For $j=1$, one might use the test statistic

$$\left| \frac{\bar{Y}_1 - \mu_0}{\sigma/\sqrt{n}} \right|$$

and see if it exceeds 1.96, i.e. testing with significance level 5%

One might similarly look at

$$\left| \frac{\bar{Y}_2 - \mu_0}{\sigma/\sqrt{n}} \right|$$

and so on.

Eventually one is going to find something "significant".

⓪

25 Sept 01

And what is the significance level of this approach?

$\text{Prob}\{\text{reject at least once} \mid H_0\}$

$= 1 - \text{Prob}\{\text{accept all} \mid H_0\}$

$= 1 - .95^J$

$J =$	1	.05
	2	.0975
	3	.143
	4	.185
	5	.226
	6	.265
	7	.302

0+

25 Sept 01
(= 24 Feb 04)

For overall level of .05

J	critical value
1	1.96
2	.63
3	.41
4	.30
5	.24
6	.20
7	.17

(Solving $1 - u^J = .95$)

0^{++}

25 Sept 01
(24 Feb 04)

Simultaneous inference

Four methods:

- Bonferroni
- Scheffé
- Tukey
- False discovery rate (FDR)

Bonferroni Inequality

Events A_j, complements $\bar{A}_j$

$$P\left(\bigcap_1^k A_j\right) \geq 1 - \sum_1^k P(\bar{A}_j)$$

$N=2$

$$P(AB) = P(A) + P(B) - P(A \cup B)$$

$$= 1 - P(\bar{A}) + 1 - P(\bar{B}) - \left(1 - P(\overline{A \cup B})\right.$$

$$= 1 - P(\bar{A}) - P(\bar{B}) + P(\overline{A \cup B})$$

$$\geq 1 - P(\bar{A}) - P(\bar{B})$$

25 Sept 01
(24 Feb 04)

O^{+++}

Bonferroni Method.

Linear combinations $\theta_j = \underset{\sim}{P}_j^T \underset{\sim}{\beta}$

$$\hat\theta_j = \underset{\sim}{P}_j' \hat{\underset{\sim}{\beta}}, \quad j=1,\dots,k$$

$$\frac{\hat\theta_j-\theta_j}{\hat\sigma\sqrt{\underset{\sim}{P}_j^T(\underset{\sim}{X}^T\underset{\sim}{X})^{-1}\underset{\sim}{P}_j}} \sim t_{n-r} \equiv \frac{\hat\theta_j-\theta_j}{\hat\sigma c_j}$$

Percent point

$$P\{ t_{n-r} \le t_{n-r}^{1-\alpha} \} = 1-\alpha$$

Bonferroni simultaneous intervals are the collection

$$\{ \hat\theta_j - \theta_j \pm t_{n-r}^{1-\alpha/2k}\, \hat\sigma c_j ; \ j=1,\dots,k \}$$

$$\text{Prob}\left\{ \frac{|\hat\theta_j-\theta_j|}{\hat\sigma c_j} \le t_{n-r}^{1-\alpha/2k}, \ j=1,\dots,k \right\}$$

$$\geq 1 - \left(\frac{\alpha}{k} + \cdots + \frac{\alpha}{k} \right)$$

$$= 1-\alpha$$ conservative

Suppose ~~things~~ of full rank 25 Sept 01

Simultaneous Confidence Intervals.

Scheffé's method.

$$\theta = \underset{k\times m}{P'}\ \underset{m\times 1}{\beta} \qquad p_1, \dots, p_k \text{ linearly independent}$$

$$\frac{(\hat\theta - \theta)^T [P^T (X^T X)^{-1} P]^{-1} (\hat\theta - \theta)/k}{\hat\sigma^2} : F_{k, m-r}$$

$$P\{ F_{k,m-r} \le F^{1-\alpha}_{k,m-r} \}$$

$$= P\{ (\hat\theta - \theta)' [\quad]^{-1} (\hat\theta - \theta) \le k \hat\sigma^2 F_{k,m-r} \}$$

$$= P\{ b' L^{-1} b \le x \} \quad \text{where} \begin{cases} b = \hat\theta - \theta \\ L = P'(X'X)^{-1}P \end{cases}$$

$$= P\left\{ \sup_{h \ne 0} \frac{(h'b)^2}{h'Lh} \le x \right\}$$

$$= P\left\{ \frac{(h'b)^2}{h'Lh} \le x, \ \forall h \ne 0 \right\}$$

$$= \text{Prob}\left\{ \frac{|h'\hat\theta - h\theta|}{\hat\sigma (h'Lh)^{1/2}} \le \left(k F^{1-\alpha}_{k,m-r}\right)^{1/2}, \ \forall h \right\}$$

Simultaneous, class of intervals

$$h'\hat\theta \pm \left(k F^{1-\alpha}_{k,m-r}\right)^{1/2} \hat\sigma \, (h'Lh)^{1/2}$$

25 Sept 0

<u>Notes</u>. 1. $\underset{\sim}{L}$ is nonsingular

$\underset{\sim}{P}'\underset{\sim}{\beta}$ estimable so $\underset{\sim}{\overset{m\times k}{P}} = \underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{w}$ for some $\underset{\sim}{\overset{m\times k}{w}}$

$$k = r(\underset{\sim}{P}) \leq r(\underset{\sim}{w}) \leq k$$

$$r(\underset{\sim}{w}) = k$$

Also var $\underset{\sim}{P}'\underset{\sim}{\hat{\beta}} = \underset{\sim}{L}\sigma^2 = \underset{\sim}{P}'(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{P}\,\sigma^2 = \underset{\sim}{w}'\underset{\sim}{w}\,\sigma^2$

<u>Lemma</u>. For $\underset{\sim}{L}$ nonsingular

$$\sup_{\underset{\sim}{h} \neq \underset{\sim}{0}} \frac{(\underset{\sim}{h}'\underset{\sim}{b})^2}{\underset{\sim}{h}'\underset{\sim}{L}\underset{\sim}{b}} = \underset{\sim}{b}'\underset{\sim}{L}^{-1}\underset{\sim}{b}$$

<u>Proof</u>.

Cauchy-Schwarz $\quad (\underset{\sim}{u}'\underset{\sim}{v})^2 \leq (\underset{\sim}{u}'\underset{\sim}{u})(v'v)$

$$v = L^{-1/2} b$$

$$u = L^{1/2} h$$

$$(h'b)^2 \leq (h'Lh)(b'L^{-1}h)$$

Back to birth example: ③ 25 Sept 01

What is leading to such a small p-value?

$\hat{\mu}_j = 7.71, 8.39, 8.52, 8.53, 8.66, 8.56, 7.71$

What is the Scheffé 5% bound?

$J = 7,\ K = 103,\ n = 721$

$r = 7,\ k' = (1\ -1)$

$L = p'(X'X)^{-1}p$

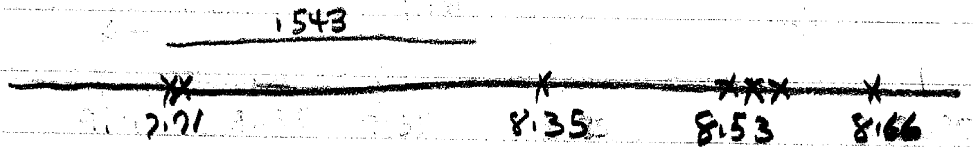
$$\bar{Y}_{j\cdot} - \bar{Y}_{j'\cdot} \pm \sqrt{(J-1) F^{.95}_{J-1,\, n-r}\, \hat{\sigma}^2 \frac{2}{K}}$$

$$1.60 \pm \sqrt{6\, F^{.95}_{6,714}\, 1.20\, \frac{2}{103}}$$

$$1.60 \pm .543$$

$F^{.05}_{6,714} = 2.11$

Because of the simultaneous nature of the confidence bounds we are allowed to examine all linear combinations and retain the overall 5% size of the test.



25 Sept. 01

Simultaneous confidence intervals for differences (Tukey)

Suppose $\bar{Y}_j - \mu_j : IN(0, \sigma^2/n)$ $\quad j = 1, \dots, J$

Tukey found q_α such that

$$\text{Prob}\left\{ \left| \frac{(\bar{Y}_i - \mu_i) - (\bar{Y}_j - \mu_j)}{s/\sqrt{n}} \right| \le q^J_{\alpha,\nu} , \forall i,j \right\} = 1 - \alpha$$

So the $J(J-1)/2$ statements

$$(\bar{Y}_i - \bar{Y}_j) - q^J_{\alpha,\nu} s/\sqrt{n} < \mu_i - \mu_j \le (\bar{Y}_i - \bar{Y}_j) + q^J_{\alpha,\nu} s/\sqrt{n}$$

are simultaneously satisfied with prob $1-\alpha$.

$q^J_{\alpha,\nu}$ is from the distribution of the studentized range

ν: df for s

Gives estimates of the differences $\mu_i - \mu_j$ plus an indication of uncertainty.

Shorter intervals than S-method. (Fewer statements)

25 Sept 01

For Tukey method need

$$q^{7}_{.05,714} = 4.170$$

Miller

(e.g. Tables in Scheffé book)

$$\bar{Y}_i - \bar{Y}_j \pm q^{7}_{\alpha,\nu} \frac{s}{\sqrt{n}}$$

$$\bar{Y}_i - \bar{Y}_j \pm 4.170 \sqrt{\frac{1.20}{103}}$$

$$\bar{Y}_i \pm \bar{Y}_j \pm .450$$

Smaller than the Scheffé limits, as was to be anticipated.