The total parameter space is 8th, 25.02.14

6

$$\Omega = \{(\mu_1, \mu_2, \dots, \mu_b, \sigma^2) : -\infty < \mu_j < \infty,\ 0 < \sigma^2 < +\infty\}$$

and

$$\omega = \{(\mu_1, \dots, \mu_b, \sigma^2) : -\infty < \mu_1 = \mu_2 = \dots = \mu_b = \mu < \infty,\ 0 < \sigma^2 < +\infty\}.$$

The likelihood functions, denoted by $L(\omega)$ and $L(\Omega)$ are, respectively

$$L(\omega) = \left(\frac{1}{2\pi\sigma^2}\right)^{ab/2} \exp\left[-\frac{1}{2\sigma^2}\sum_{j=1}^{b}\sum_{i=1}^{a}(x_{ij}-\mu)^2\right]$$

and

$$L(\Omega) = \left(\frac{1}{2\pi\sigma^2}\right)^{ab/2} \exp\left[-\frac{1}{2\sigma^2}\sum_{j=1}^{b}\sum_{i=1}^{a}(x_{ij}-\mu_j)^2\right].$$

Now

$$\frac{\partial \log L(\omega)}{\partial \mu} = \sigma^{-2}\sum_{j=1}^{b}\sum_{i=1}^{a}(x_{ij}-\mu)$$

and

$$\frac{\partial \log L(\omega)}{\partial (\sigma^2)} = -\frac{ab}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{j=1}^{b}\sum_{i=1}^{a}(x_{ij}-\mu)^2.$$

Solving $\frac{\partial \log L(\omega)}{\partial \mu} = 0$ and $\frac{\partial \log L(\omega)}{\partial (\sigma^2)} = 0$, we obtain

$$\hat{\mu} = \bar{x}_{..} = \frac{1}{ab}\sum_{j=1}^{b}\sum_{i=1}^{a} x_{ij},$$

$$\hat{\sigma}_0^2 = v = \frac{1}{ab}\sum_{j=1}^{b}\sum_{i=1}^{a}(x_{ij}-\bar{x}_{..})^2,$$

Sufficient condition!

and these values maximize $L(\omega)$.

Furthermore,

$$\frac{\partial \log L(\Omega)}{\partial \mu_j} = \sigma^{-2} \sum_{i=1}^{a} (x_{ij} - \mu_j), \quad j = 1, 2, \dots, b,$$

and
$$\frac{\partial \log L(\Omega)}{\partial (\sigma^2)} = -\frac{ab}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \mu_j)^2$$

~~Then~~ As a result

$$\hat{\mu}_j = \bar{x}_{.j} = \frac{1}{a} \sum_{i=1}^{a} x_{ij}, \quad j = 1, 2, \dots, b$$

sufficient condition?

$$\hat{\sigma}_1^2 = w = \frac{1}{ab} \sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \bar{x}_{.j})^2$$

maximize $L(\Omega)$. These maxima are, respectively,

$$L(\hat{\omega}) = \left[\frac{ab}{2\pi \sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \bar{x}_{..})^2} \right]^{ab/2} \exp\left[- \frac{ab \sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \bar{x}_{..})^2}{2 \sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \bar{x}_{..})^2} \right]$$

$$= \left[\frac{ab}{2\pi \sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \bar{x}_{..})^2} \right]^{ab/2} \exp\left[-\frac{ab}{2} \right]$$

and

$$L(\hat{\Omega}) = \left[\frac{ab}{2\pi \sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \bar{x}_{.j})^2} \right]^{ab/2} \exp\left[-\frac{ab}{2} \right].$$

Finally,

$$\Lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})} = \left[\frac{\sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \bar{x}_{.j})^2}{\sum_{j=1}^{b} \sum_{i=1}^{a} (x_{ij} - \bar{x}_{..})^2} \right]^{ab/2} := \left[\frac{Q_3}{Q} \right]^{\frac{ab}{2}}.$$

We reject the hypothesis H_0 if $\Lambda \leqslant \lambda_0$.

We find λ_0.

we have

$$\frac{Q_3}{Q} = \frac{Q_3}{Q_3 + Q_4} = \frac{1}{1 + \frac{Q_4}{Q_3}}.$$

~~The significance level of the test of H_0 is~~ Therefore,

$$\alpha = P_{H_0}\left[\frac{1}{1 + Q_4/Q_3} \leq \lambda_0^{\frac{2}{ab}}\right] \overset{!}{=} P_{H_0}\left[\frac{\frac{Q_4}{b-1}}{\frac{Q_3}{b(a-1)}} \geq c\right],$$

where

$$c = \frac{b(a-1)}{b-1}\left(\lambda_0^{-2/ab} - 1\right).$$

But

$$F = \frac{\frac{Q_4}{\sigma^2(b-1)}}{\frac{Q_3}{\sigma^2 b(a-1)}} = \frac{\frac{Q_4}{b-1}}{\frac{Q_3}{b(a-1)}}$$

has an F-distribution with $b-1$ and $b(a-1)$ degrees of freedom.

As a result, ~~The constant~~ c ~~is so selected as to yield the desired value of α i.e~~

$$c = q_{F(b-1, b(a-1))}(1-\alpha).$$

Remark 2

The samples may be of different sizes, for instance,

$a_1, a_2, \ldots, a_b$

5. The Analysis of Variance

Recall that the one-way analysis of variance (ANOVA) problem ~~considered~~ concerned one factor at b levels. Now, we have two factors A and B with levels a and b, respectively.

Let X_{ij}, $i = 1, .., a$ and $j = 1, .., b$ denote the response for Factor A at level i and Factor B at level j.

Denote the total sample size by $n = ab$. We shall assume that the $X_{ij}'s$ are iid $N(\mu_{ij}, \sigma^2)$.

The mean μ_{ij} is often referred to as the mean of the (i,j)th cell.

First, we will consider the additive model, where

$$\mu_{ij} = \bar{\mu} + (\bar{\mu}_{i.} - \bar{\mu}) + (\bar{\mu}_{.j} - \bar{\mu});$$

that is, the mean in the (i,j)th cell is due to additive effects of the levels, i of Factor A and j of Factor B, over the average (constant) $\bar{\mu}$.

Let $\alpha_i = \bar{\mu}_{i\cdot} - \bar{\mu}$, $i = 1, \dots, a$; $\beta_j = \bar{\mu}_{\cdot j} - \bar{\mu}$, $j = 1, \dots, b$; and $\mu = \bar{\mu}$. Then

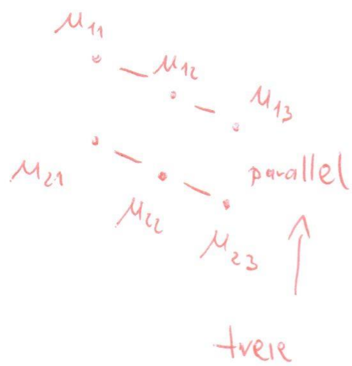
$$\mu_{ij} = \mu + \alpha_i + \beta_j,$$

where $\sum_{i=1}^{a} \alpha_i = 0$ and $\sum_{j=1}^{b} \beta_j = 0$. We refer to this model as be~~i~~ng a <u>two-way</u> ANOVA model.

<u>Example 1 (Mean profile plots)</u>

For $a = 2$, $b = 3$, $\mu = 5$, $\alpha_1 = 1$, $\alpha_2 = -1$, $\beta_1 = 1$, $\beta_2 = 0$, $\beta_3 = -1$, we have

		Factor B		
		1	2	3
Factor A	1	$\mu_{11} = 7$	$\mu_{12} = 6$	$\mu_{13} = 5$
	2	$\mu_{21} = 5$	$\mu_{22} = 4$	$\mu_{23} = 3$



If $\beta_1 = \beta_2 = \beta_3 = 0$, we have

		Factor B		
		1	2	3
Factor A	1	$\mu_{11} = 6$	$\mu_{12} = 6$	$\mu_{13} = 6$
	2	$\mu_{21} = 4$	$\mu_{22} = 4$	$\mu_{23} = 4$

The hypotheses of interest are,

$$H_{0A}: \alpha_1 = \dots = \alpha_a = 0 \quad \text{versus} \quad H_{1A}: \alpha_i \neq 0, \text{ for some } i,$$

and

$$H_{0B}: \beta_1 = \dots = \beta_b = 0 \quad \text{versus} \quad H_{1B}: \beta_j \neq 0, \text{ for some } j.$$

If H_{0A} (H_{0B}) is true then the mean of the (i,j)th cell does not depend on the level of A (B). [cf. example, case 2]

We call these hypotheses <u>main effect</u> hypotheses.

The likelihood ratio test for H_{0B} versus H_{1B}

Recall that $Q = Q_4 + Q_3$. That is

$$\underbrace{(ab-1)S^2}_{\substack{\text{total sum of squares}\\ [TSS]\\ [SST]}} = \underbrace{\sum_{j=1}^{b}\sum_{i=1}^{a}(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot})^2}_{\text{sum of squares among columns means}} + \underbrace{\sum_{j=1}^{b}\sum_{i=1}^{a}(X_{ij} - \bar{X}_{\cdot j})^2}_{\text{sum of squares within columns}}$$

is decomposed into

needless

we have shown it on exercises. 499

Recall that $Q = Q_2 + Q_4 + Q_5$. That is

$$\underbrace{(ab-1)s^2}_{TSS} = \underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2}_{\text{sum of squares among rows}} + \underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot})^2}_{\text{sum of square among columns}} + \underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b}(X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X}_{\cdot\cdot})^2}_{\text{sum of squares the remains}}.$$

is decomposed

$Q/\sigma^2 \sim \chi^2(ab-1)$

$\frac{Q_2}{\sigma^2} \sim \chi^2(a-1)$

$\frac{Q_4}{\sigma^2} \sim \chi^2(b-1)$

$\frac{Q_5}{\sigma^2} \sim \chi^2([a-1][b-1])$

(LRT for H_{0B} against H_{1B})

The ~~A~~ test statistic ~~has a form~~

$$F = \frac{\frac{Q_4}{b-1}}{\frac{Q_5}{(a-1)(b-1)}}$$

has, under H_{0B}, an F-distribution with $b-1$ and $(a-1)(b-1)$ degrees of freedom.

The hypothesis H_{0B} is rejected if $F \geq c$, where $\alpha = P_{H_{0B}}(F \geq c)$.

We shall compute ~~a the~~ the distribution of F under ~~the~~ the an alternatives.

We have

$$E[X_{ij}] = \mu + \alpha_i + \beta_j$$

$$E[\bar{X}_{i\cdot}] = \mu + \alpha_i, \quad E[\bar{X}_{\cdot j}] = \mu + \beta_j, \text{ and } E[\bar{X}_{\cdot\cdot}] = \mu.$$

The noncentrality parameter Q_4/σ^2 is

$$\frac{a}{\sigma^2}\sum_{j=1}^{b}(\mu + \beta_j - \mu)^2 = \frac{a}{\sigma^2}\sum_{j=1}^{b}\beta_j^2$$

and that of Q_5/σ^2 is

$$\frac{1}{\sigma^2}\sum^{b}\sum^{a}(\mu + \alpha_i + \beta_j - \mu - \alpha_i - \mu - \beta_j + \mu)^2 = 0$$

Thus, if the hypothesis H_{1B} is true, F has a noncentral F-distribution with $b-1$ and $(a-1)(b-1)$ degrees of freedom and noncentrality parameter $\frac{a}{\sigma^2}\sum_{j=1}^{b}\beta_j^2$.

The likelihood ratio test for H_{0A} versus H_{1A}

The test statistic

$$F = \frac{\frac{Q_2}{a-1}}{\frac{Q_5}{[a-1][b-1]}}$$

has, under H_{0A}, an F-distribution with $a-1$ and $(a-1)(b-1)$ degrees of freedom. The hypothesis H_{0A} is rejected if $F \geq c$, where $\alpha = P_{H_{0A}}(F \geq c)$.

If the hypothesis H_{1A} is true, F has a noncentral F-distribution with $a-1$ and $(a-1)(b-1)$ degrees of freedom and noncentrality parameter $\frac{b}{\sigma^2}\sum_{i=1}^{a}\alpha_i^2$.

Remark 1

The above analysis-of-variance problem is usally referred to as a two-way classification with one observation per cell.

Let X_{ijk}, $i=1,..,a$; $j=1,..,b$; $k=1,..,c$, (denote) $n=abc$ random variables which are independent and have normal distributions with common, but unknown, variance σ^2. Denote the mean of each X_{ijk}, $k=1,..,c$ by μ_{ij}.

Consider the parameters,

$$\gamma_{ij} = \mu_{ij} - \{\mu + (\bar{\mu}_{i.} - \mu) + (\bar{\mu}_{.j} - \mu)\}, \quad i=1,..,a, \; j=1,..,b.$$
$$= \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \mu$$

These parameters are called interaction parameters.

~~and~~ [They reflect the specific contribution to the cell mean over and above the additive model]