1.a. $y = \alpha + \beta(x-\bar{x}) \Rightarrow \begin{cases} \hat{\alpha} = \bar{y} \\ \hat{\beta} = \sum y_i(x_i-\bar{x}) / \sum(x_i-\bar{x})^2 \end{cases} \Rightarrow \begin{cases} \hat{\alpha} = 0.4078 \\ \hat{\beta} = 0.0657 \end{cases}$

$y_0 = \hat{\alpha} + \hat{\beta}(x_0 - \bar{x}) \Rightarrow \hat{x}_0 = \frac{y_0 - \hat{\alpha}}{\hat{\beta}} + \bar{x} = 9.1996$

CI for $\hat{x}_0$ equivalent to CI for $\theta = \frac{y_0 - \hat{\alpha}}{\hat{\beta}} \triangleq \frac{u}{v}$, $u \perp v$ b.c. $\hat{\alpha} \perp \hat{\beta}$

$Var(u) = var(y_0) + var(\hat{\alpha}) = \sigma^2(1+\frac{1}{n}) \triangleq a_1\sigma^2$

$Var(v) = var(\hat{\beta}) = \sigma^2\left(\frac{1}{\sum(x_i-\bar{x})^2}\right) \triangleq a_2\sigma^2$

using Fieller's Theorem, $(1-\alpha)\%$ CI for $\theta = \frac{u}{v}$ is: $\frac{1}{(1-g)}\left(\frac{u}{v} \pm \frac{t_{n-2,\alpha/2} \cdot s}{v}\sqrt{(1-g)a_1^2 + \frac{u^2 a_2^2}{v^2}}\right)$

where $s^2 = \frac{1}{(n-2)}(y_i - \hat{y})^2$, $g = \frac{t^2_{n-2,\alpha/2} \cdot s^2 \cdot a_2^2}{v^2}$

plug in $\alpha = 0.05$, $n = 8$, and the $\{(x_i, y_i)\}$ data:

95% CI for $\theta = \frac{y_0 - \hat{\alpha}}{\hat{\beta}}$ is $(1.5116, 3.6378)$

thus: 95% CI for $\hat{x}_0 = \frac{y_0 - \hat{\alpha}}{\hat{\beta}} + \bar{x} = \theta + \bar{x}$ is $(8.1366, 10.2628)$

1.b. $EY = X\beta$, $X \in R^{n\times p}$, $rk(X) = p$. $cov(Y) = \sigma^2 I_n$

$\hat{\beta} = (X'X)^{-1}X'Y$, $cov(\hat{\beta}) = \sigma^2 \cdot (X'X)^{-1}$, $E(\hat{\beta}) = \beta$.

let $\hat{\theta} = \frac{c'\hat{\beta}}{d'\hat{\beta}}$, since $E(c'\hat{\beta}) = c'\beta$, $E(d'\hat{\beta}) = d'\beta$,

$\triangleq \frac{u}{v}$

$var(c'\hat{\beta}) = c' cov(\hat{\beta}) c = \sigma^2 \cdot c'(X'X)^{-1}c \triangleq v_{11}\sigma^2$

$var(d'\hat{\beta}) = d' cov(\hat{\beta}) d = \sigma^2 d'(X'X)^{-1}d \triangleq v_{22}\sigma^2$

$cov(c'\hat{\beta}, d'\hat{\beta}) = c' cov(\hat{\beta}) d = \sigma^2 c'(X'X)^{-1}d \triangleq v_{12}\sigma^2$

from Fieller's Theorem: $100(1-\alpha)\%$ CI for $\theta = \frac{c'\beta}{d'\beta}$ is:

$$\frac{1}{(1-g)}\left[\frac{u}{v} - \frac{g v_{12}}{v_{22}} \pm \frac{t_{r,\alpha/2} \cdot s}{v}\left(v_{11} - 2\frac{u}{v}v_{12} + \frac{u^2}{v^2}v_{22} - g\left(v_{11} - \frac{v_{12}^2}{v_{22}}\right)\right)^{1/2}\right]$$

where $r = n-p$; $s^2 = \frac{1}{n-p}\sum(y_i - \hat{y}_i)^2$, $\hat{y}_i$ is the predicted value from linear regression ;

$g = \frac{t^2_{r,\alpha/2} \cdot s^2 \cdot v_{22}}{v^2}$; $u = c'\hat{\beta}$, $v = d'\hat{\beta}$;

$v_{11} = c'(X'X)^{-1}c$, $v_{22} = d'(X'X)^{-1}d$, $v_{12} = c'(X'X)^{-1}d$.

1.c. $y_{ij} = \mu + a_i + \varepsilon_{ij}$, $i = 1, 2, \dots k$, $j = 1, 2, \dots n_k$, $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_i^2)$

let $z = (\bar{y}_{1\cdot}, \bar{y}_{2\cdot}, \dots \bar{y}_{k\cdot})^T$, $\bar{y}_{i\cdot} \sim N(\mu + a_i, \sigma_k^2/n_k)$, $\bar{y}_{i\cdot} \perp \bar{y}_{j\cdot}$, $\forall i \neq j$

$z \sim N(\mu_z, \Sigma_z)$, $\mu_z = (\mu + a_1, \dots, \mu + a_k)^T$, $\Sigma_z = \operatorname{diag}(\sigma_1^2/n_1, \dots \sigma_k^2/n_k)$

denote $\vec{a} = \begin{pmatrix} \mu + a_1 \\ \vdots \\ \mu + a_k \end{pmatrix}$. Testing linear combination of group means equal to zero

$\Leftrightarrow H_0: \theta = c'a = 0, \vec{c} = (c_1, \dots, c_k)^T$ is a known vector

let $\hat{\theta} = c'z$, then $\hat{\theta} \sim N(\mu_\theta, \delta_\theta)$ where $\mu_\theta = E(\hat{\theta}) = c'E(z) = c'a$

$$\delta_\theta = \operatorname{cov}(\hat{\theta}) = c'\Sigma_z c = \sum_{i=1}^k \frac{c_i^2}{n_i} \cdot \sigma_i^2$$

denote $s_i^2 = \frac{1}{n_i - 1}\left(\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2\right)$, $u_i \triangleq \frac{s_i^2}{\sigma_i^2} \cdot (n_i - 1) \sim \chi^2(n_i - 1)$, $u_i \perp\!\!\!\perp \bar{y}_{j\cdot}$, $\forall i, j \in [1,]$

test statistic
$$t = \frac{(c'z - c'a)}{\left(\sum_{i=1}^k \frac{c_i^2}{n_i} \cdot s_i^2\right)^{1/2} / b^{1/2}} \;\dot\sim\; t_b.$$
$$= \frac{(c'z - c'a) / \left(\sum_{i=1}^k \frac{c_i^2}{n_i} \cdot \sigma_i^2\right)}{\frac{1}{b^{1/2}}\left(\frac{\sum c_i^2/n_i \cdot s_i^2}{\sum c_i^2/n_i \cdot \sigma_i^2}\right)^{1/2}} \triangleq \frac{z}{\sqrt{X_b / b}}$$

since $z \sim N(0, 1)$, we only need to approximate distribution of X_b, and corresponding $df = b$.

$$X_b = \sum_{i=1}^k \left(\sum c_i^2/n_i \cdot \sigma_i^2\right)^{-1} \cdot \frac{c_i^2}{n_i} \cdot s_i^2$$
$$= \sum_{i=1}^k \left(\sum c_i^2/n_i \cdot \sigma_i^2\right)^{-1} \cdot \frac{\sigma_i^2 c_i^2}{n_i(n_i - 1)} \cdot u_i$$
$$\triangleq \sum_{i=1}^k g_i \cdot u_i, \quad u_i \overset{\text{indept}}{\sim} \chi^2(n_i - 1)$$

with Satterthwaite approximation: $\frac{b \cdot X_b}{E(X_b)} \sim \chi^2_b$ when

$$\hat{b} = \left(\sum_{i=1}^k g_i u_i\right)^2 \Big/ \sum_{i=1}^k \frac{(g_i u_i)^2}{df_i}$$
$$= \frac{\left(\sum c_i^2/n_i \cdot \sigma_i^2\right)^2}{\left(\sum c_i^2/n_i \cdot \sigma_i^2\right)^2} \cdot \left(\sum_{i=1}^k c_i^2 \cdot \frac{s_i^2}{n_i}\right)^2 \Big/ \sum_{i=1}^k \left(c_i^2 \frac{s_i^2}{n_i}\right)^2 / (n_i - 1)$$

denote $w_i = s_i^2 / n_i$, then $\hat{b} = \left(\sum_{i=1}^k c_i^2 w_i\right)^2 \Big/ \sum_{i=1}^k \left(\frac{c_i^4 w_i^2}{n_i - 1}\right)$

using the b estimated by $\hat{b}$ shown above:

$$t = \frac{c'z}{\sqrt{\left(\sum_{i=1}^k c_i^2 \frac{s_i^2}{n_i}\right) / b}} \;\dot\sim\; t_b,$$

given confidence level of α, we can reject $H_0: c'a = 0$ when $|t| > t_{b, 1-\alpha/2}$, upper $\alpha/2$ quantile of t_b.

2.9. $y = X\beta + e$, $E(e) = 0$, $E(e'e) = cov(e) = \sigma^2 I_n$, $b = (X'X)^{-1}X'Y$

leverage of i th case: $h_{ii} = \{X(X'X)^{-1}X'\}_{(i,i)}$

i th internal residual: $r_i = \dfrac{e_i}{s\sqrt{1-h_{ii}}}$, $s^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \hat{\sigma}^2$

with ith: $cov(b) = (X'X)^{-1}X' \, cov(Y) \, X(X'X)^{-1}$

$$= \sigma^2 \cdot (X'X)^{-1}$$

without ith: $X_{(i)}$: X without ith row; $Y_{(i)}$: Y without ith row.; $S_{(i)}^2$, $\hat{\sigma}^2$ without ith case.

$$cov(b_{(i)}) = (X'_{(i)}X_{(i)})^{-1}X'_{(i)} \, cov(Y_{(i)}) \, X_{(i)}(X'_{(i)}X_{(i)})^{-1}$$

$$= \sigma^2 \cdot (X'_{(i)}X_{(i)})^{-1}$$

$$VR_i = \frac{|\widehat{cov}(b)|}{|\widehat{cov}(b_{(i)})|} = \frac{|s^2 \cdot X'X|^{-1}}{|S_{(i)}^2 X'_{(i)}X_{(i)}|^{-1}} = \left(\frac{S_{(i)}^2}{S^2}\right)^p \frac{|X'_{(i)}X_{(i)}|}{|X'X|}$$

① $(n-p-1)S_{(i)}^2 = (n-p)s^2 - \dfrac{e_i^2}{1-h_{ii}}$

$$\frac{s^2}{S_{(i)}^2} = \frac{(n-p-1)s^2}{(n-p)s^2 - e_i^2/1-h_{ii}} = \frac{n-p-1}{n-p-r_i^2}$$

② let $X' = (\vec{x}_1, \dots \vec{x}_n)$, then $X'X = \sum_{j=1}^{n} x_j x_j'$

$$X'_{(i)}X_{(i)} = \sum_{j=1, j\neq i}^{n} x_j x_j' = X'X - x_i x_i'$$

$$\Rightarrow \det(X'_{(i)}X_{(i)}) = \det(X'X - x_i x_i')$$

$$= \det(X'X)(1 - x_i'(X'X)^{-1}x_i)$$

$$= \det(X'X)(1 - h_{ii})$$

$$\Rightarrow \frac{|X'_{(i)}X_{(i)}|}{|X'X|} = 1 - h_{ii}$$

from ①.② $VR_i = \left(\dfrac{S_{(i)}^2}{S^2}\right)^p \cdot \dfrac{|X'_{(i)}X_{(i)}|}{|X'X|}$

$$= \left(\frac{n-p-r_i^2}{n-p-1}\right)^p \cdot (1-h_{ii})$$

interpretation:

$|\widehat{cov}(b)|$, also called the genaralized variance, is a scalar measure for the overall precision of the β estimates. A smaller $|cov(b)|$ indicates a more precise estimation, and vice versa.

$VR_i = \frac{|\widehat{cov}(b)|}{|\widehat{cov}(b_i)|} > 1$: precision is degraded by the ith observation

$VR_i < 1$: overall precision is improved by the ith observation.

26. Bartlett's test

Bartlett's test is used to test if observations from k different samples have a same variance. Whenever we assume that outcomes follow normal distribution with same variance across groups or samples, we can use Bartlett test to verify it.

The construction of test statistic for Bartlett's test is based on the principle that all exact tests of composite hypotheses are equivalent to tests of simple hypotheses for conditional samples. That is, to test variation in sample S that is independent of irrelevant unknowns ϕ, we should construct a set of sufficient set of statistics u for ϕ. Then all the exact tests of significance independent of ϕ must be and will only be based on the conditional distr. of $S|u$.

In Bartlett's test, we want to test for the "goodness of fit" involving all the remaining degrees of freedom of the sample, then the test should be based only on the conditional likelihood of $T|u$. $T = (S_1^2, \ldots S_k^2)$, the within group sample variances, and $u = s^2$, pooled residual variance, sufficient for the irrelevant parameters.

All we need to do is to find the distribution of a function of $T|u$, $G(T|u)$, and the we reject Ho: within group variances don't individually fit better than overall variance i.e. $H_0: \sigma_1^2 = \ldots = \sigma_k^2$, when the statistic is extreme enough.

For any statistic of form s^2 (sample variance df $= n$), in normal theory, $L(s^2)$ must be invariant for a simultaneous change in scale in both s^2 and σ^2. We can derive:

$$L(s^2) = \text{Const.} + n\log\left(\frac{s}{\sigma}\right) - \frac{n}{2}\left(\frac{s^2}{\sigma^2} - 1\right)$$

Then we can define a likelihood-ratio-type statistic as:

$\Lambda = -2\log\mu = -2\log(f(T|u)+\text{const}) = (\sum_{i=1}^{k} n_i)\log s^2 - \sum_{i=1}^{k} n_i \log s_i^2$, n_i is the df of s_i^2

From construction of T.S, we know that μ has a limiting normal form. Then

$\Lambda = -2\log\mu$ can be approximated by a χ^2 distr. with an unknown df.

charenteristic

$$M_\Lambda(t) = E(\mu)^{-2t} = \Gamma\{\tfrac{n}{2}(1-2t)\}(\tfrac{n}{2})^{nt} e^{-nt} \Big/ (1-2t)^{\frac{n}{2}}\Gamma(\tfrac{n}{2})$$

$$K(n) = \log M_\Lambda(t) = t(1+\tfrac{1}{3n}+\cdots) + \tfrac{t^2}{2}(2+\tfrac{4}{3n}+\cdots) + \tfrac{t^2}{3!}(8+\tfrac{24}{3n}+\cdots)+\cdots$$

$$= t(1+\tfrac{1}{3n}) + 2\tfrac{t^2}{2!}(1+\tfrac{1}{3n})^2 + 8\tfrac{t^3}{3!}(1+\tfrac{1}{3n})^3 \quad ①$$

$$K = \sum_{i=1}^{k} K(n_i) - K(n)$$

$$M_{\chi^2_q}(t) = (1-2t)^{-k/2}$$

$$K_{\chi^2_q} = \log M_{\chi^2_q}(t) = -\tfrac{k}{2}\cdot\log(1-2t)$$

$$= -\tfrac{q}{2}(-2t - \tfrac{2^2}{2}t^2 - \tfrac{2^3}{3}t^3 - \cdots)$$

$$= qt + qt^2 + \tfrac{4}{3}qt^3 + \cdots \quad ②$$

comparing ①, ②. $q = k-1$, and negelect $O(\frac{1}{n_i^2})$'s

$$\Rightarrow \frac{-2\log\mu}{c} \sim \chi^2_{k-1}, \quad c = 1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{n_i} - \frac{1}{\sum_{i=1}^{k} n_i}\right)$$

To summarize, $G(T|u) = -\dfrac{2\log\mu}{c} = \dfrac{(\sum_{i=1}^{k} m_i - k)\ln s^2 - \sum_{i=1}^{k}(m_i-1)\ln s_i^2}{1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{m_i-1} - \frac{1}{\sum_{i=1}^{k} m_i - k}\right)} \sim \chi^2_{k-1}$.

where $m_i = n_i + 1$. is the # of observation in group i.

We can reject $H_0: \sigma_1^2 = \cdots = \sigma_k^2$ when $G(T|u) > \chi^{2\,1-\alpha}_{k-1}$.

In the public health framework, the Bartlett's test can be used to compare the effect stability among different batches / treatments. For example. if we conduct 1-way ANOVA for the treatment effect of three medications A, B, C, and there are no significant different in mean effects. Then we can conduct Bartlett's test for equal variance. If we reject H_0, then we know that although 3 medication seems to have same effects, at least two of them have different stability. (or consistency)

2.C.

In regression model. $Y = X\beta + \varepsilon$. $Y \in R^{n\times 1}$, $X \in R^{n\times p}$, $\varepsilon \in R^{n\times 1}$. $rk(X) = p$

$\varepsilon \sim N(0, \sigma^2 \cdot diag(w_1, \dots w_n))$

To test if variance depend on some predictors, we can parameterize it as:

Testing if there is a function $w()$ s.t. $w_i = w(\vec{z}_i, \vec{\lambda})$, $\exists i_1 \neq i_2, w_{i_1} \neq w_{i_2}$

where. $\vec{z}_i = (x_{i,j_1}, x_{i,j_2}, \dots x_{i,j_q})^T \in R^{q\times 1}$, $1 \le j_1 < \dots < j_q \le p$

$\vec{\lambda} = (\lambda_1, \dots \lambda_k)^T \in R^{k\times 1}$, some additional parameters.

To conduct a score test mention in Cook and Weisburg's paper 'Diagnostics for heteroscedasity in regression'. We furtherly assume that $w(\vec{z}, \vec{\lambda})$ follows, $\exists \vec{\lambda}_0$. s.t. $w(\vec{z}_1, \vec{\lambda}_0) = \dots = w(\vec{z}_n, \vec{\lambda}_0)$. (e.g. $w(\vec{z}, \vec{\lambda}) = \exp(\vec{z}'\lambda)$)

With all assumptions above met, and an explicit form of $w(\vec{z}, \vec{\lambda})$, we can conduct the score test for $H_0: w_1 = \dots = w_n$, i.e., $H_0: \vec{\lambda} = \vec{\lambda}_0$ as following:

Step 1: $e = (I - X(X'X)^{-1}X')Y$, $\hat{\sigma}^2 = \frac{1}{n} e'e$, $u = \begin{pmatrix} e_1^2 \\ \vdots \\ e_n^2 \end{pmatrix} \frac{1}{\hat{\sigma}^2}$

Step 2: $w'(\vec{z}, \vec{\lambda}) = \dfrac{\partial w(\vec{z}, \vec{\lambda})}{\partial \vec{\lambda}} = \begin{pmatrix} \dfrac{\partial w(\vec{z}, \vec{\lambda})}{\partial \lambda_1} \\ \vdots \\ \dfrac{\partial w(\vec{z}, \vec{\lambda})}{\partial \lambda_k} \end{pmatrix}$

Step 3: $D = \begin{bmatrix} [w'(\vec{z}_1, \vec{\lambda}_0)]^T \\ \vdots \\ [w'(\vec{z}_n, \vec{\lambda}_0)]^T \end{bmatrix}$, $\bar{D} = D - 11'D/n$

Step 4: explicitly. $S = \frac{1}{2} u'\bar{D}(\bar{D}'\bar{D})^{-1}\bar{D}'u$

or computationally: Regression u on $[1_n | D]$. and let $S = \frac{1}{2}$(sum of square)

Step 5: compare S with $\chi^2(q)$ (q: length of $\vec{z}_i$)

Given significance level $= \alpha$:

If $S > \chi^{2\,1-\alpha}_q$, reject H_0, variance of outcome depend on predictors included in z_i's

If $S \le \chi^{2\,1-\alpha}_q$, cannot reject H_0. variance of outcome doesn't depend on z_i's

3. $Y_{ij} = \mu + a_i + \varepsilon_{ij}$, $i = 1, 2, \dots b$. $j = 1, 2, \dots n$

$a_i \overset{iid}{\sim} N(0, \sigma_a^2)$, $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$, $a \perp \varepsilon$

(a) $SSTr = n \sum_{i=1}^{b} (a_i - \bar{\varepsilon}_{i\cdot} - (\bar{a} - \bar{\varepsilon}))^2$

as $a_i - \bar{\varepsilon}_{i\cdot} \sim N(0, \sigma_a^2 + \sigma^2/n)$, $(\bar{a} - \bar{\varepsilon})$ = sample mean of $(a_i - \bar{\varepsilon}_{i\cdot})$

$SSTr = (n\sigma_a^2 + \sigma^2) \underbrace{\sum_{b=1}^{b} \left[\frac{a_i - \bar{a} + \bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}}{\sigma_a^2 + \sigma^2/n} \right]^2} \sim \chi^2(b-1)$ (sample variance of $a_1 - \bar{\varepsilon}_1., \dots a_b - \bar{\varepsilon}_b$

$\Rightarrow SSTr/(n\sigma_a^2 + \sigma^2) \sim \chi^2(b-1)$

$\Rightarrow E(SSTr) = (b-1)(n\sigma_a^2 + \sigma^2)$

$SSE = \sum_{i=1}^{b} \sum_{j=1}^{n} (\varepsilon_{ij} - \bar{\varepsilon}_{i\cdot})^2$

$= \sigma^2 \cdot \sum_{i=1}^{b} \left[\underbrace{\sum_{j=1}^{n} \left(\frac{\varepsilon_{ij} - \bar{\varepsilon}_{i\cdot}}{\sigma} \right)^2} \right] \sim \chi^2(n-1)$, sample variance of $\varepsilon_{i1}, \dots \varepsilon_{in}$

$\Rightarrow SSE/\sigma^2 \sim \chi^2(b(n-1))$

$\Rightarrow E(SSE) = b(n-1)\sigma^2$

then : $\frac{SSTr/(n\sigma_a^2 + \sigma^2)}{SSE/\sigma^2} \cdot \frac{b(n-1)}{(b-1)} \sim F_{b-1, b(n-1)}$

$\frac{\sigma^2}{n\sigma_a^2 + \sigma^2} \cdot \frac{MSTr}{MSE} \sim F_{b-1, b(n-1)}$

$\theta = \frac{\sigma_a^2}{\sigma^2}$ $\qquad \frac{1}{n\theta + 1} \cdot \frac{MSTr}{MSE} \sim F_{b-1, b(n-1)}$

$\Rightarrow$ 95% CI for θ is $\left\{ \theta : F_{b-1, b(n-1)}^{.025} \le \frac{MSTr}{MSE} \frac{1}{n\theta + 1} \le F_{b-1, b(n-1)}^{.975} \right\}$

$\Rightarrow$ 95% CI for θ is : $\left(\frac{1}{F_{b-1, b(n-1)}^{.975}} \cdot \frac{MSTr}{MSE} - 1 \right) \cdot \frac{1}{n} \le \theta \le \left(\frac{1}{F_{b-1, b(n-1)}^{.025}} \cdot \frac{MSTr}{MSE} - 1 \right) \cdot \frac{1}{n}$.

(b) power $= P(f > F \mid H_1)$

$H_0: \sigma_a^2 = 0$, H_1, $\sigma_a^2 > 1$, $\sigma_a^2 = C_0 \sigma^2$ (for power derivation)

as $\frac{(b-1) MSTr}{n\sigma_a^2 + \sigma^2} \sim \chi^2(b-1)$, $\frac{(n-1) b MSE}{\sigma^2} \sim \chi^2(n-1)b$

under $H_0: \sigma_a^2 = 0$, $f = \frac{MSTr}{MSE} \sim F_{b-1, bn-b}$

given α, reject $\sigma_a^2 = 0$ when $f > F_{b-1, bn-b}^{1-\alpha}$, upper α quantile

under $H_1: \sigma_a^2 = C_0 \sigma^2$ on the other hand

$g = \frac{MSTr/(n\sigma_a^2 + \sigma^2)}{MSE/\sigma^2} \sim F_{b-1, bn-b}$

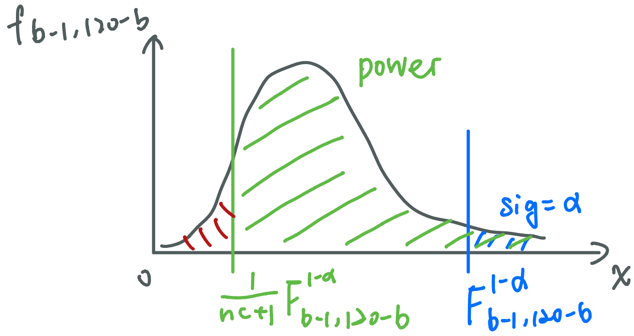
$$q = f \cdot \frac{\sigma^2}{n\sigma_a^2 + \sigma^2} = f \cdot \frac{1}{nc+1} \sim F_{b-1,\, bn-b}$$

Therefore: Power $= P(\text{reject } H_0 \mid H_1)$

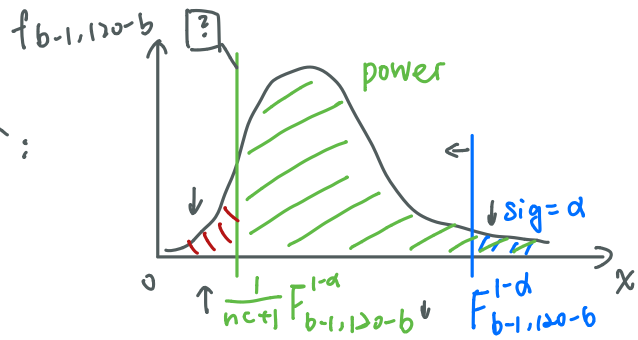
$$= P\left(f > F^{1-\alpha}_{b-1,\, bn-b} \;\middle|\; f \cdot \frac{1}{nc+1} \sim F_{b-1,\, bn-b}\right)$$

$$= 1 - \Phi\left(F^{1-\alpha}_{b-1,\, bn-b} \cdot \frac{1}{nc+1}\right), \quad \Phi \text{ is the cdf of } F_{b-1,\, bn-b}$$

$$= 1 - \Phi\left(F^{1-\alpha}_{b-1,\, 120-b} \cdot \frac{1}{nc+1}\right)$$



as $b \uparrow$:



The power depends on position of $q = \frac{1}{nc+1} F^{1-\alpha}_{b-1,\, 120-b}$, and the shape of pdf of F-distribution, $df1 = b-1$, $df2 = 120-b$.

To simplify the problem, and only consider the effect of $df1$, $df2$, denote $q = \text{const.} \cdot F^{1-\alpha}_{b-1,\, 120-b}$, $\text{const.} \ll 1$.

Consider the pdf of $F_{df1,\, df2}$, a larger $df1$ or a smaller $df2$ yields lighter tail and lighter starting (smaller green and red area).

Thus smaller $F^{1-\alpha}_{df1,\, df2}$, smaller q, and larger power.

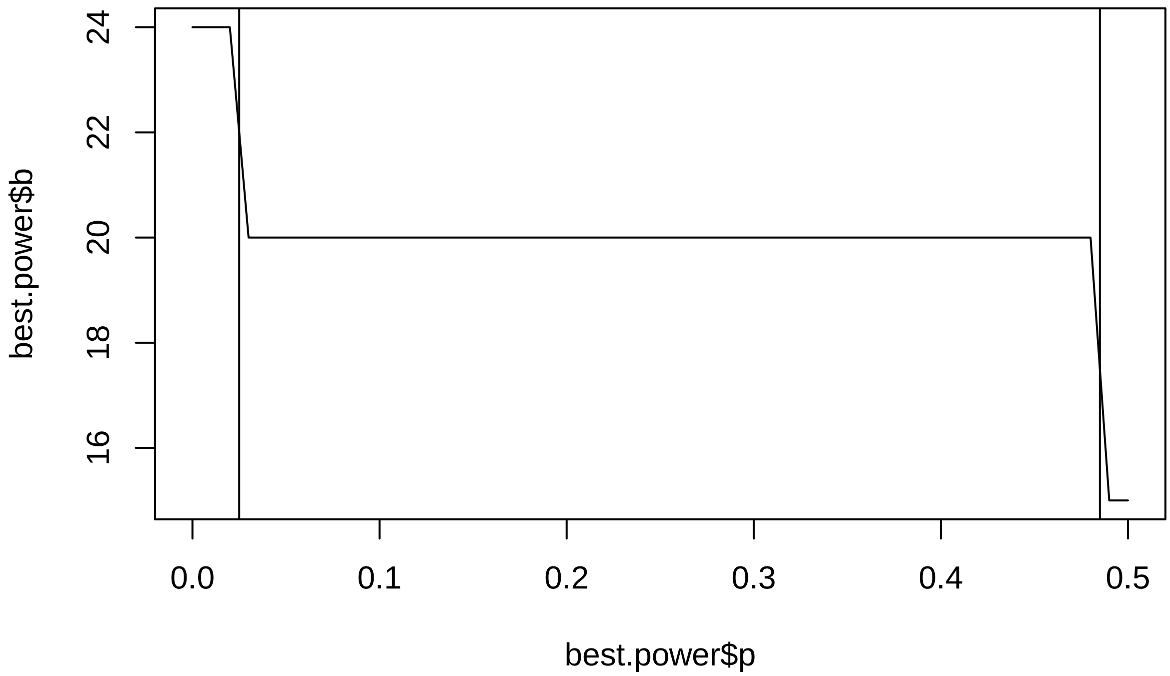
(However, $q = \frac{1}{nc+1} F^{1-\alpha}_{b-1,\, 120-b} = \frac{1}{120c/b+1} \cdot F^{1-\alpha}_{b-1,\, 120-b}$, as $b \uparrow$, $\frac{1}{120c/b+1} \uparrow$, $F^{1-\alpha}_{b-1,\, 120-b} \downarrow$

therefore, with the constraint of $nb = 120$, the power is not monotone with b).

3.c

```r
# candidates for b and n
b <- c(2,3,4,5,6,8,10,12,15,20,24,30,40,60)
n <- 120/b
# c = sigma_a / sigma
c <- c(1/0.5, 1/1, 1/1.5)
# power results for each combination
power <- matrix(0, 14, 3)
for (i in 1:3){
  c.temp <- c[i]
  for (j in 1:14){
    b.temp <- b[j]
    n.temp <- n[j]
    q <- qf(0.99, b.temp - 1, 120 - b.temp) / (n.temp * c.temp + 1)
    power[j,i] <- 1 - pf(q, b.temp - 1, 120 - b.temp)
  }
}
```

```r
# grid for p(sigma = 0.5) = p(sigma = 1.5) = p
p.seq <- seq(0, 0.5, 0.01)
best.power <- matrix(rep(p.seq, 3), ncol = 3, byrow = FALSE)
for (i in 1:length(seq(0, 0.5, 0.01))){
  p <- p.seq[i]
  q <- 1- 2 * p
  # given a set of (b,n), sigma follows discrete distribution
  # E(power) = sum (power_i * p(sigma = sigma_i))
  temp <- (power[,1] + power[,3]) * p + power[,2] * q
  # for a given p, select the best power and the best (b,n)
  best.power[i, 2] <- max(temp)
  best.power[i, 3] <- which.max(temp)
}
best.power <- data.frame(best.power)
colnames(best.power) <- c("p", "power", "index")
best.power$b <- b[best.power$index]
best.power$n <- n[best.power$index]
```



- When $p < 0.03$, $(b,n) = (24,5)$ will maximize the power.
- When $0.03 <= p < 0.49$, $(b,n) = (20, 6)$ will maximize the power.
- When $p >= 0.49$, $(b,n) = (15, 8)$ will maximize the power.

4.a. $y_{ijk} = \mu + \underline{\tau_i} + \underline{\beta_j} + (\tau\beta)_{ij} + e_{ijk}$, $i=1,2,\dots a$, $j=1,2,\dots b$, $k=1,2,\dots n$

A: fixed B: random

$e_{ijk} \overset{iid}{\sim} N(0,\sigma^2)$, $\beta_j \overset{iid}{\sim} N(0,\sigma_\beta^2)$, $\tau\beta_{ij} \sim N(0, \frac{a-1}{a}\sigma_{\tau\beta}^2)$

$\sum_{i=1}^a \tau_i = 0$, $\sum_{i=1}^a \tau\beta_{ij} = 0$, $\forall j=1,2,\dots b$

Covariance:

for fixed j:

let $TB_j = \begin{pmatrix} \tau\beta_{1j} \\ \vdots \\ \tau\beta_{aj} \end{pmatrix}$, suppose $TB_j \sim N(0, cI_a)$, should be indept. without constraint

$$z = \begin{pmatrix} TB_j \\ 1^T TB_j \end{pmatrix} = \begin{pmatrix} I_a \\ 1^T \end{pmatrix} TB_j, \quad cov(z) = cov\begin{pmatrix} TB_j \\ 1^T TB_j \end{pmatrix} = \begin{pmatrix} cI_a & c\cdot 1 \\ c\cdot 1' & ca \end{pmatrix}$$

$$\Rightarrow (TB_j \mid 1^T TB_j = 0) \sim N(0, cI_a - 11^T \frac{c^2}{ca})$$

since $\begin{cases} TB_{ij} \sim N(0, (\frac{a-1}{a})\sigma_{\tau\beta}^2) \\ \sum_{i=1}^a TB_{ij} = 0 \end{cases}$ $\Rightarrow \{cI_a - 11^T \frac{c}{a}\}_{jj} = (\frac{a-1}{a})\sigma_{\tau\beta}^2$

$\Leftrightarrow c(\frac{a-1}{a}) = \sigma_{\tau\beta}^2 \Leftrightarrow c = \sigma_{\tau\beta}^2$

Therefore: $TB_j = \begin{pmatrix} \tau\beta_{1j} \\ \vdots \\ \tau\beta_{aj} \end{pmatrix} \sim N(0, \sigma_{\tau\beta}^2 I_a)$

and $TB_j \mid 1' TB_j = 0 \sim N(0, \sigma_{\tau\beta}^2 (I_a - 11^T \frac{1}{a}))$, $\forall j=1,2,\dots b$

that is: given $\sum_{i=1}^a \tau\beta_{ij} = 0$ for $\forall j=1,2\dots b$.

$$cov(\tau\beta_{ij}, \tau\beta_{lk}) = \begin{cases} (\frac{a-1}{a})\sigma_{\tau\beta}^2, & i=l, j=k \\ -\frac{1}{a}\sigma_{\tau\beta}^2, & i\neq l, j=k \\ 0, & j\neq k \end{cases}$$

Point estimation:

$$SSB = \sum_k \sum_i \sum_{j=1}^b (\beta_j - \bar\beta + \bar\varepsilon_{\cdot j \cdot} - \bar\varepsilon)^2$$

$$E(SSB) = an\, E\Big[\underbrace{\sum_{j=1}^b \big((\beta_j + \bar\varepsilon_{\cdot j \cdot}) - \overline{(\beta_j + \bar\varepsilon_{\cdot j \cdot})}\big)^2}_{}\Big]$$

$(b-1)$(sample var of $\beta_j + \bar\varepsilon_{\cdot j \cdot}$) $\sim (\sigma_\beta^2 + \frac{1}{na}\sigma^2)\cdot \chi^2(b-1)$

$$= (an\sigma_\beta^2 + \sigma^2)(b-1)$$

$$SSAB = \sum_k \sum_{i=1}^a \sum_{j=1}^b (\bar y_{\cdots} + \bar y_{ij\cdot} - \bar y_{i\cdot\cdot} - \bar y_{\cdot j \cdot})^2$$

$$= \sum_k \sum_{i=1}^a \sum_{j=1}^b (\underline{\mu} + \underline{\bar\beta} + \bar\varepsilon_{\cdots} + \underline{\mu} + \underline{\tau_i} + \underline{\beta_j} + \tau\beta_{ij} + \bar\varepsilon_{ij\cdot} - \underline{\mu} - \underline{\tau_i} - \underline{\bar\beta} - \overline{\tau\beta}_{i\cdot} - \bar\varepsilon_{i\cdot\cdot} - \underline{\mu} - \underline{\beta_j} - \bar\varepsilon_{\cdot j \cdot})^2$$

$$= \sum_k \sum_{i=1}^{a} \sum_{j=1}^{b} \left(\tau\beta_{ij} - \overline{\tau\beta}_{i\cdot} + \bar{\varepsilon}_{\cdots} + \bar{\varepsilon}_{ij\cdot} - \bar{\varepsilon}_{i\cdot\cdot} - \bar{\varepsilon}_{\cdot j\cdot} \right)^2$$

$$= \sum_k \sum_{j=1}^{b} \sum_{i=1}^{a} \left[(\bar{\varepsilon}_{\cdots} + \bar{\varepsilon}_{ij\cdot} - \bar{\varepsilon}_{i\cdot\cdot} - \bar{\varepsilon}_{\cdot j\cdot})^2 + (\tau\beta_{ij} - \overline{\tau\beta}_{i\cdot})^2 \right] + 0$$

$$E(SSAB) = E\left[\sum_k \sum_{j=1}^{b} \sum_{i=1}^{a} \left((\bar{\varepsilon}_{ij\cdot} - \bar{\varepsilon}_{\cdot j\cdot}) - (\bar{\varepsilon}_{\cdots} - \bar{\varepsilon}_{i\cdot\cdot}) \right)^2 \right] + nE\left[\sum_{j=1}^{b} \sum_{i=1}^{a} (\tau\beta_{ij} - \overline{\tau\beta}_{i\cdot})^2 \right]$$

$$= (a-1)(b-1)(n\sigma^2_{\tau\beta} + \sigma^2)$$

$$SSE = \sum_k \sum_j \sum_i \left[y_{ijk} - \bar{y}_{ij\cdot} \right]^2 = \sum_j \sum_i \sum_{k=1}^{n} \left[\varepsilon_{ijk} - \bar{\varepsilon}_{ij\cdot} \right]^2$$

$$E(SSE) = ab(n-1)\sigma^2$$

$$\Rightarrow \hat{\sigma}^2 = \frac{SSE}{ab(n-1)} = MSE$$

$$\hat{\sigma}^2_\beta = \frac{1}{an}\left(\frac{SSB}{b-1} - \frac{SSE}{ab(n-1)} \right) = \frac{MSB - MSE}{an}$$

$$\hat{\sigma}^2_{\tau\beta} = \frac{1}{n}\left(\frac{SSAB}{(a-1)(b-1)} - \frac{SSE}{ab(n-1)} \right) = \frac{MSAB - MSE}{n}$$

<u>test</u> : $H_0: \sigma^2_\beta = 0$, under H_0, $\frac{SSB}{SSE} \sim F_{b-1, ab(n-1)}$

therefore, given significance level $= \alpha$,

reject H_0 when $\frac{SSB}{SSE} > F^{1-\alpha}_{b-1, ab(n-1)}$

4.6. $y_{ijk} = \mu + \alpha_i + \pi_j + (\alpha\pi)_{ij} + \varepsilon_{ijk}$, $i = 1, \dots a$, $j = 1, \dots b$, $k = 1, \dots n$.

$\pi \perp\!\!\!\perp \alpha\pi \perp\!\!\!\perp \varepsilon$

$$SSB = \sum_k \sum_i \sum_j (\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})^2 = \sum_k \sum_i \sum_j \left(\pi_j - \bar{\pi} + \overline{(\alpha\pi)}_{\cdot j} - \overline{(\alpha\pi)}_{\cdot\cdot} + \bar{\varepsilon}_{\cdot j\cdot} - \bar{\varepsilon}_{\cdots} \right)^2$$

$$E(SSB) = a \cdot n \cdot E\left[\sum_{j=1}^{b} \left(\pi_j + \overline{(\alpha\pi)}_{\cdot j} + \bar{\varepsilon}_{\cdot j\cdot} - \overline{(\pi_j + \overline{(\alpha\pi)}_{\cdot j} + \bar{\varepsilon}_{\cdot j\cdot})} \right)^2 \right]$$

$\sim (\sigma^2_\pi + \frac{1}{a}\sigma^2_{\pi\alpha} + \frac{1}{an}\sigma^2) \cdot \chi^2_{(b-1)}$

$$= an(b-1)(\sigma^2_\pi + \frac{1}{a}\sigma^2_{\pi\alpha} + \frac{1}{an}\sigma^2)$$

$$SSA = \sum_k \sum_j \sum_i (\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2 = \sum_k \sum_j \sum_i \left(\alpha_i - \bar{\alpha} + \overline{(\alpha\pi)}_{i\cdot} - \overline{(\alpha\pi)}_{\cdot\cdot} + \bar{\varepsilon}_{i\cdot\cdot} - \bar{\varepsilon}_{\cdots} \right)^2$$

$$= \sum_k \sum_j \left[\sum_i (\alpha_i - \bar{\alpha})^2 + \sum_i \left[\overline{(\alpha\pi)}_{i\cdot} - \overline{(\alpha\pi)}_{\cdot\cdot} + \bar{\varepsilon}_{i\cdot\cdot} - \bar{\varepsilon}_{\cdots} \right]^2 \right]$$

$$E(SSA) = bn \sum_i \alpha_i^2 + bn(a-1)\left(\frac{\sigma^2_{\alpha\pi}}{b} + \frac{\sigma^2}{bn} \right)$$

$$SSAB = \sum_k \sum_j \sum_i (\bar{y}_{\cdots} + \bar{y}_{ij\cdot} - \bar{y}_{\cdot j\cdot} - \bar{y}_{i\cdot\cdot})^2$$

$$= n \cdot \sum_j \sum_i \Big(\bar{\pi} + \overline{\alpha\pi}_{\cdot\cdot} + \bar{\varepsilon}_{\cdots} + \alpha_i + \pi_j + \alpha\pi_{ij} + \bar{\varepsilon}_{ij\cdot}$$
$$- \pi_j - \overline{\alpha\pi}_{\cdot j} - \bar{\varepsilon}_{\cdot j\cdot} - \alpha_i - \bar{\pi} - \overline{\alpha\pi}_{i\cdot} - \bar{\varepsilon}_{i\cdot\cdot} \Big)^2$$

$$= n \sum_j \sum_i \left(\alpha\pi_{ij} - \overline{\alpha\pi}_{\cdot j} + \bar{\varepsilon}_{ij\cdot} - \bar{\varepsilon}_{\cdot j\cdot} + \overline{\alpha\pi}_{\cdot\cdot} - \overline{\alpha\pi}_{i\cdot} + \bar{\varepsilon}_{\cdots} - \bar{\varepsilon}_{i\cdot\cdot} \right)^2$$

$$E(SSAB) = n E \sum_j \sum_i \left[(\alpha\pi_{ij} - \overline{\alpha\pi}_{\cdot j})^2 + (\bar{\varepsilon}_{ij\cdot} - \bar{\varepsilon}_{\cdot j\cdot})^2 + (\overline{\alpha\pi}_{\cdot\cdot} - \overline{\alpha\pi}_{i\cdot})^2 + (\bar{\varepsilon}_{\cdots} - \bar{\varepsilon}_{i\cdot\cdot})^2 \right]$$

$$= n \left[ab\,\sigma^2_{\alpha\pi} - \frac{ab}{a} \sigma^2_{\alpha\pi} + \sigma^2_{\alpha\pi} - \frac{ab}{b} \sigma^2_{\alpha\pi} + \frac{\sigma^2}{n} (ab - a - b + 1) \right)$$

$$= (a-1)(b-1)(n\,\sigma^2_{\alpha\pi} + \sigma^2)$$

$$SSE = \sum_k \sum_i \sum_j (\varepsilon_{ijk} - \bar{\varepsilon}_{ij\cdot})^2$$

$$E(SSE) = ab(n-1)\sigma^2$$

(i)

df	term	A f i	B r j	ERROR r k	E(MS)
$a-1$	α_i	0	b	n	$bn \frac{\sum \pi_i^2}{a-1} + n\sigma^2_{\pi a} + \sigma^2$
$b-1$	π_j	a	1	n	$an\sigma^2_{\pi} + n\sigma^2_{\pi a} + \sigma^2$
$(a-1)(b-1)$	$\alpha\pi_{ij}$	1	1	n	$n\sigma^2_{\pi a} + \sigma^2$
$ab(n-1)$	ε_{ijk}	1	1	1	σ^2

Test for $H_0: \sigma^2_\pi = 0$, under H_0, $\frac{MSB}{MSAB} \sim F_{b-1,(a-1)(b-1)}$

thus, given α, reject $H_0: \sigma^2_\pi = 0$, when $\frac{MSB}{MSAB} > F^{1-\alpha}_{b-1,(a-1)(b-1)}$.

(ii) No, this is not the same as (a).

4.c. Although we changed the model set-up, the data was still the same. Therefore, the E(MS)'s stays the same, only that we used them to estimate different terms.

$$\Rightarrow \hat{\sigma}^2_\beta = \frac{MSB - MSE}{an}, \quad \hat{\sigma}^2_\pi = \frac{MSB - MSAB}{an}, \quad \hat{\sigma}^2_{\alpha\pi} = \frac{MSAB - MSE}{n}$$

$$\Rightarrow \hat{\sigma}^2_\beta = \hat{\sigma}^2_\pi - \frac{1}{a} \hat{\sigma}^2_{\alpha\pi}$$

Application: As in model with constraint (a), and the model without (b), the estimation of variance of random effects can be expressed by each other, we can always add constraint on interaction, and take advantage of (a)'s computational simplicity, for both point estimation and F-test