Analysis of Covariance.

(1) Introduction:

It's generalization of ANOVA models that the design matrix contains both qualitative and quantitative explanatory variables.

Rmk: Qualitative variable are: regression variables. covariates. concomitant variables.

Two goals:

i) Compare treatments

ii) Inference on regression coefficient correspond covariates

Rmk: Concomitant variables are intended to serve as blocking factors to sharpen the analysis — access the difference of treatments — reduce the variability.

So ANCOVA can be viewed as a variance reduction design.

Two applications:

i) Missing data

ii) BIBD.

e.g. Consider $Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma x_{ijk} + \varepsilon_{ijk}$.

α_i. β_j are qualitative factors. x_{ijk} is the concomitant variable. γ is regression coefficient

Next, we will consider $Y = (X\ Z)\begin{pmatrix}\beta \\ \gamma\end{pmatrix} + \varepsilon$

$X \in M^{n\times p}$, $Z \in M^{n\times s}$, $\varepsilon \sim N_n(0, \sigma^2 I_n)$

Rmk: Z is introduced to sharpen analysis. its coefficient γ are tested after the ANOVA tests.

① Estimation of γ:

· Written the model in: $Y = X\beta + MZ\gamma + (I-M)Z\gamma + \varepsilon$

i.e. $Y = X\delta + (I-M)Z\gamma + \varepsilon$. $\delta = \beta + (X^TX)^{-}X^TZ\gamma$.

$$\Rightarrow \begin{pmatrix} (X^TX)\delta \\ Z^T(I-M)Z\gamma \end{pmatrix} = \begin{pmatrix} X^TY \\ Z^T(I-M)Y \end{pmatrix} \text{ (Normal Equation)}$$

i) $Z^T(I-M)Z$ is nonsingular:

Then γ is estimable. $\hat{\gamma} = (Z^T(I-M)Z)^{-1}Z^T(I-M)Y$

so $X\beta$ is estimable. $\widehat{X\beta} = MY - MZ\hat{\gamma}$

But commonly. β isn't estimable.

ii) $Z^T(I-M)Z$ is singular:

Then γ. $X\beta$ are not estimable. generally.

Besides, $E(X\hat{\beta}) = E(MY - MZ\hat{\gamma})$

$$= X\beta + MZ\gamma - MZ(Z^T(I-M)Z)^- Z^T(I-M)Z\gamma$$

Next, we characterize estimable function of γ:

<u>Thm.</u> $\xi^T\gamma$ is estimable $\Leftrightarrow \exists \ell \in \mathbb{R}^n$. $\xi^T = \ell^T(I-M)Z$

<u>Pf:</u> ($\Rightarrow$) $\xi^T\gamma = \ell^T(X\ Z)\begin{pmatrix}\beta\\ \gamma\end{pmatrix}$ $\therefore \ell^T X = 0$

$$\ell^T Z = \ell^T(I-M)Z.$$

($\Leftarrow$) $\ell^T(I-M)Z\gamma = \ell^T(I-M)(X\ Z)\begin{pmatrix}\beta\\ \gamma\end{pmatrix}$

<u>Cr.</u> $Z\gamma$ is estimable $\Leftrightarrow$ $C(Z) \cap C^\perp(X) = C(Z)$. i.e. $MZ=0$

② <u>Estimation of σ^2:</u>

If $Z^T(I-M)Z$ is nonsingular. Then $(X\beta, \gamma)$ are both estimable. $\operatorname{cov}\begin{pmatrix}\widehat{X\beta}\\ \hat{\gamma}\end{pmatrix} = \sigma^2\begin{pmatrix}A_{11} & A_{12}\\ A_{12}^T & A_{22}\end{pmatrix}$

$A_{11} = M(I + Z(Z^T(I-M)Z)^{-1}Z^TM)$ $\quad A_{22} = (Z^T(I-M)Z)^{-1}$

$A_{12} = -MZ(Z^T(I-M)Z)^{-1}$

Generally. set $P = M_{(X,Z)}$. $SSE = \|(I-P)Y\|^2$.

Note that $C(X\ Z) = C(X\ (I-M)Z)$

$\Rightarrow P = X(X^TX)^-X^T + (I-M)Z(Z^T(I-M)Z)^-Z^T(I-M)$

<u>Denote:</u> $E_{AB} = A^T(I-M)B$.

$\Rightarrow Y^T(I-P)Y = E_{yy} - E_{yz}E_{zz}^-E_{zy}$.

③ Hypothesis Testing:

The primary interest in test is treatment effect.

For $C(X_0) \subset C(X)$. We want to reduce:

$Y = X_0\gamma_0 + Z\gamma + \varepsilon$ from $\Rightarrow Y = X\beta + Z\gamma + \varepsilon$.

i.e. $H_0: E(Y) \in C(X_0, Z)$ v.s. $H_1: E(Y) \in C(X_0, Z)^c \cap C(X, Z)$

$$F = \frac{\|(P-P_0)Y\|^2 / r(P-P_0)}{\|(I-P)Y\|^2 / r(I-P)} \sim F(r(P-P_0), r(I-P), \gamma^*)$$

$$\gamma^* = \frac{\|(P-P_0)(X\beta + Z\gamma)\|^2}{\sigma^2}, \quad \gamma^* \overset{H_0}{=} 0.$$

$P_0 = M(X_0, Z) = X_0(X_0^T X_0)^- X_0^T + (I-M_0)Z(Z^T(I-M_0)Z)^- Z^T(I-M_0)$

Rmk: $Y^T(P-P_0)Y = Y^T(I-P_0)Y - Y^T(I-P)Y$

$= Y^T\tilde{M}_0 Y - Y^T\tilde{M}Y.$

④ SS for general ANCOVA:

i) Consider $Y = X\beta + Z\gamma + \varepsilon$. where the ANOVA part is two-way balanced ANOVA with interaction.

$n = abN$. total observation.

$C(X) = C(M_\mu) + C(M_\alpha) + C(M_\eta) + C(M_{\alpha\eta})$

For testing. e.g. $H_0: (\alpha\eta)_{11} = \cdots = (\alpha\eta)_{ab}$

Then $M_0 = M - M_{\alpha\eta}$.

For testing concomitant γ: $H_0: \gamma = 0$ v.s. $H_1: \gamma \neq 0$.

i.e. $H_0: E(Y) \in C(X)$ v.s. $H_1: E(Y) \in C(X\ Z)$.

$P_0 = M. \Rightarrow \|(P-P_0)Y\|^2 = E_{\eta z} E_{zz}^{-1} E_{z\eta}$.

ii) Consider balanced two-way ANOVA with no replication and one covariate. Then ANCOVA model:

$Y_{ij} = \mu + \alpha_i + \eta_j + \gamma Z_{ij} + \varepsilon_{ij}$. $1 \le i \le a$. $1 \le j \le b$. $n = abN$. $N = 1$.

$E_{\eta\eta} = Y^T(I-M)Y = \sum_{i=1}^{a}\sum_{j=1}^{b}(Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2$

$E_{\eta z}$. E_{zz} are analogous.

To test $H_0: \gamma = 0$. $\Rightarrow P_0 = M$ $r(P_0) = a+b-1$

$$\therefore F = \frac{\|(P-P_0)Y\|^2 / r(P-P_0)}{\|(I-P)Y\|^2 / r(I-P)} = \frac{E_{\eta z} E_{zz}^{-1} E_{z\eta}}{(E_{\eta\eta} - E_{\eta z} E_{zz}^{-1} E_{z\eta})/(n-a-b)}$$

$$\sim F(1, n-a-b, \gamma^*)$$

(2) Application:

① Missing Data:

Suppose some responses are missing from model:

$Y = (y_1 \cdots y_n)^T$ is n responces. $Y = X\beta + \varepsilon$ is the complete model. $Y = (Y_{n-r}\ Y_r)^T$. If the last r components Y_r is missing. The model become:

$Y_{n-r} = X_{n-r}\beta + \varepsilon_{n-r}$. $X = \begin{pmatrix} X_{n-r} \\ X_r \end{pmatrix}$. $\cdots$ (Δ).

$\Rightarrow$ Introduce covariate vectors $Z_i = (0 \cdots \overset{i^{th}}{1}\, 0 \cdots 0)^T_{n\times 1}$

for each missing data Y_i.

<u>Denote</u>: $Z = \begin{pmatrix} 0 \\ I_r \end{pmatrix}$. $\tilde{Y} = \begin{pmatrix} Y_{n-r} \\ 0 \end{pmatrix}$

written the model into ANCOVA model:

$$\tilde{Y} = \begin{pmatrix} X_{n-r} \\ X_r \end{pmatrix} \beta + \begin{pmatrix} 0 \\ I_r \end{pmatrix} \gamma + \varepsilon = (X\ Z) \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \varepsilon \cdots (*)$$

i) <u>SSE of $(*)$</u>:

$$C(X\ Z) = C \begin{pmatrix} X_{n-r} & 0 \\ 0 & I_r \end{pmatrix} \Rightarrow P = \begin{pmatrix} M_{n-r} & 0 \\ 0 & I_r \end{pmatrix}$$

$\Rightarrow \tilde{Y}^T (I-P) \tilde{Y} = Y_{n-r}^T (I - M_{n-r}) Y_{n-r}$. SSE of $(*)$

It's identical with SSE of (Δ).

ii) <u>Estimation of β of $(*)$</u>:

Note that $E(\tilde{Y}) = \begin{pmatrix} X_{n-r}\beta \\ X_r \beta + \gamma \end{pmatrix} = (X\ Z) \begin{pmatrix} \beta \\ \gamma \end{pmatrix}$

Estimation: $\widehat{E(\tilde{Y})} = P\tilde{Y} = \begin{pmatrix} M_{n-r} Y_{n-r} \\ 0 \end{pmatrix}$. So it

is identical with $X\hat{\beta} = M_{n-r} Y_{n-r}$ in (Δ).

<u>Prop</u>. Estimable Function of β in (Δ) ~~is~~ is estimable in $(*)$. And their estimates are identical.

<u>Pf</u>: The estimable function of (Δ) has form:

$\ell_{n-r}^T X_{n-r} \beta$.

For $(*)$ is: $\ell^T (X\beta + Z\gamma)$

There're equal $\iff$ $\ell^T = (\ell_{n-r}\ 0)$. i.e. $\ell^T Z = 0$

Then: $\ell^T P \tilde{Y} = (\ell_{n-r}\ 0)\begin{pmatrix} M_{n-r} & 0 \\ 0 & I_r \end{pmatrix}\begin{pmatrix} Y_{n-r} \\ 0 \end{pmatrix}$

$= \ell_{n-r} M_{n-r} Y_{n-r} = \widehat{\ell_{n-r} X_{n-r}\beta}$ in (Δ).

iii) Estimation of Missing Values:

1°) Assume y is estimable. Then $\hat{y} = (Z^T(I-M)Z)^{-1} Z^T(I-M)\tilde{Y}$

2°) Construct complete data of responses $Y^* = \tilde{Y} - Z\hat{y}$

3°) Use the data Y^* to fit complete data model (∇)

Rmk: Partition $M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{pmatrix}$. $Z = \begin{pmatrix} 0 \\ I_r \end{pmatrix}$. Then:

Simplify: $\hat{y} = (I_r - M_{22})^{-1} M_{21} Y_{n-r}$.

And it's linear function of Y_{n-r}, which retains full information of X design matrix. can be seen as the prediction of Y_r by Y_{n-r}, X.

4°) For complete data model $Y^* = X\beta + \varepsilon$ (∇)

SSE in (∇) is identical in $(*)$.

$Y^{*T}(I-M)Y^* = (\tilde{Y}^T - (Z\hat{y})^T)(I-M)(\tilde{Y} - Z\hat{y})$

$= \tilde{Y}^T(I-M)\tilde{Y} - \tilde{Y}^T P_{((I-M)Z)} \tilde{Y}$

$= \tilde{Y}^T(I-P)\tilde{Y} = Y_{n-r}^T(I-M_{n-r})Y_{n-r}$

estimate of $X\beta$ in (∇) is identical in $(*)$.

$\widehat{X\beta}$ in (∇) $= MY^* = M(\tilde{Y} - Z\hat{y}) = \widehat{X\beta}$ in $(*)$

So does $\mathrm{Var}(\ell^T\widehat{X\beta})$. for $\ell \in \mathbb{R}^n$.

$$Var(c^T X\hat{\beta}) = Var(c^T(M - MZ(Z^T(I-M)Z)^{-1}Z^T(I-M)\tilde{Y}))$$
$$= \sigma^2(c^TMc + c^TMZ(Z^T(I-M)Z)^{-1}Z^TMc)$$

Rmk: i) It's not simply $\sigma^2 c^T M c$, follows from the result of missing data.

ii) In sum. estimate of estimable Func's of β in (Δ) is identical in $(*)$, (∇)

② Balanced Incomplete Block Design (BIBD):

· Suppose we set b blocks, t treatments. The number of treatments can be observed in each block is k $(k < t)$

BIBD is a design that each pair of treatments occur together in the fixed block a fix times: λ.

Set r is number of replications for each treatment.

So: $tr = bk$. Denote it by n. total number.

Besides, the pairs containing a fixed treatment occur $(t-1)\lambda = r(k-1)$ times

Rmk: i) The two condition implies: $b \geq t$.

ii) The conditions are neccessary but not sufficient condition for BIB exists.

e.g. $b = t = 4$. $\lambda = 2$, $k = 3$. B_i's are blocks.

B_1	B_2	B_3	B_4
A	A	A	\
B	B	\	B
C	\	C	C
\	D	D	D

BIB model can be written as:

$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$. $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$. $i = 1, 2, \cdots b$.

$j \in D_i$ the set of indices of treatments in block i

Denote: A_j is set of indices of block where the treatment j occurs.

Rmk: i) $|D_i| = k$. $|A_j| = r$

ii) Written the model in ANCOVA:

$$Y = (X\ Z)\begin{pmatrix}\beta\\ \tau\end{pmatrix} + \varepsilon.\ \beta = \begin{pmatrix}\mu\\ \beta_1\\ \vdots\\ \beta_b\end{pmatrix}.\ \tau = \begin{pmatrix}\tau_1\\ \vdots\\ \tau_t\end{pmatrix}$$

$X \in M^{bk \times (b+1)}$, $Z \in M^{bk \times t}$, β_i is block effect. τ_j is treatment effect.

Our primary interest is treatment effects.

Note that to compute: $\hat{\tau} = (Z^T(I-M)Z)^{-1} Z^T(I-M)Y$.

Write $Z = (Z_1 \cdots Z_t)$. $Z_m = [Z_{ij,m}]$ $Z_{ij,m} = \delta_{jm}$.

i.e. $Z_m = 0$ for all rows except r rows equal 1.

To Compute $Z^T(I-M)Z = Z^TZ - Z^TMZ$

i) Find $Z_m^T Z_s$:

1') $m = s$: $Z_m^T Z_m = \sum_{j=1}^{t} \sum_{i \in A_j} \delta_{jm} = r$

2') $m \neq s$: $Z_m^T Z_s = \sum_{j=1}^{t} \sum_{i \in A_j} \delta_{jm} \delta_{js} = 0$

$\Rightarrow Z^T Z = r I_t$

ii) <u>Find $Z^T M Z$:</u>

$M = (V_{ij,i'j'})$ $\quad V_{ij,i'j'} = \frac{1}{k} \delta_{ii'}$

Denote: $M Z_m = (d_{ij,m})$

$$d_{ij,m} = \sum_{i',j'} V_{ij,i'j'} Z_{i'j',m} = \sum_{j'=1}^{t} \sum_{i' \in A_{j'}} \frac{1}{k} \delta_{ii'} \delta_{j'm}$$

$$= \sum_{i' \in A_m} \frac{1}{k} \delta_{ii'} = \frac{1}{k} \delta_i (A_m)$$

i.e. If treatment m is in block i. Then. $d_{ij,m} = 1/k$.

$$\Rightarrow Z_m^T M Z_m = \sum_{i=1}^{b} \sum_{j \in D_i} k^{-2} \delta_i (A_m) = \frac{r}{k}$$

$$Z_s^T M Z_m = \sum_{i=1}^{b} \sum_{j \in D_i} \frac{1}{k^2} \delta_i (A_s) \delta_i (A_m) = \frac{\lambda}{k}$$

Thus. $Z^T M Z = \begin{pmatrix} \frac{\lambda}{k} & \frac{r}{k} & \cdots & \frac{r}{k} \\ \frac{r}{k} & \ddots & & \vdots \\ \vdots & & \ddots & \\ \frac{r}{k} & \cdots & & \frac{\lambda}{k} \end{pmatrix} = \frac{1}{k} ((r-\lambda) I + \lambda J_t^t)$

iii) <u>Summary:</u>

$$Z^T (I - M) Z = r I - \frac{1}{k} ((r-\lambda) I + \lambda J_t^t)$$

$$= \frac{1}{k} ((r(k-1) + \lambda) I - \lambda J_t^t)$$

Note that $r(k-1) = (t-1)\lambda$. Denote $W = I - \frac{1}{t} J_t^t$

$\Rightarrow Z^T (I - M) Z = \frac{\lambda t}{k} W$. (W is orthonormal proj.)

$$\Rightarrow (Z^T(I-M)Z)^- = \frac{k}{\lambda t} W^- = \frac{k}{\lambda t} W$$

With : $Y^T(I-M)Z_m = \sum_{ij} (\eta_{ij} - \bar{\eta}_{i\cdot}) Z_{ij.m}$

$$= \sum_{i \in A_m} (\eta_{ij} - \bar{\eta}_{i}) =: Q_m$$

$$\Rightarrow Y^T(I-M)Z = (Q_1 \cdots Q_t)$$

iv) Estimation of τ :

Now we get estimable func. of τ : $\xi^T\hat{\tau}$, $\xi^T = e^T(I-M)Z$.

i.e. $\xi^T\hat{\tau} = e^T(I-M)Z(Z^T(I-M)Z)^- Z^T(I-M)Y$.

$Var(\xi^T\hat{\tau}) = \sigma^2 \xi^T (Z^T(I-M)Z)^- \xi = \frac{k\sigma^2}{\lambda t} \xi^T\xi$.

Rmk: $(I-M)Z(Z^T(I-M)Z)^-$ can be simplified :

$$ZJ_t = J_n \quad (I-M)J_n = 0 \text{ i.e. } (I-M)ZJ_t = 0$$

$$(I-M)Z(Z^T(I-M)Z)^- = (I-M)Z\frac{k}{\lambda t}(I - \tfrac{1}{t}J_tJ_t^T)$$

$$= \frac{k}{\lambda t}(I-M)Z$$

$$\Rightarrow \xi^T\hat{\tau} = \frac{k}{\lambda t}\xi^T(Q_1 \cdots Q_t)^T = \frac{k}{\lambda t}\sum_{j=1}^{t} \xi_j Q_j$$

v) Test of $\xi^T\tau$:

$$SSE = Y^T(I - M - (I-M)Z(Z^T(I-M)Z)^- Z^T(I-M))Y$$

$$= Y^T(I-M)Y - \frac{k}{\lambda t}Y^T(I-M)ZZ^T(I-M)Y$$

$$= \sum(\eta_{ij} - \bar{\eta}_{i\cdot})^2 - \frac{k}{\lambda t}\sum^{t} Q_j^2$$

$$\Rightarrow MSE = SSE / (bk - t - b + 1)$$

To test $H_0 : \xi^T\tau = 0$.

Then $F = \dfrac{(\xi^T\hat{z})^2 / \frac{k}{\lambda t}\xi^T\xi}{MSE} \sim F(1, bk-b-t+1, \gamma)$

$\gamma = (\xi^T z)^2 / \sigma^2 \xi^T\xi \frac{k}{\lambda t}$. $\quad \gamma \overset{H_0}{=} 0$.

Rmk: For block part: $Y^T(M - P_n)Y = \frac{1}{k}\sum_1^b B_i^2 - \frac{G^2}{n}$.

where $B_i = \sum_{j \in D_i} Y_{ij}$. $G = \sum_{ij} Y_{ij}$

③ Incomplete Blocks Design:

In general IBD model, any treatment arrangement is permissible. e.g.

B_1	B_2	B_3
A	C	E
B	D	`
A	`	`