STA305 Midterm Review

- $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
- randomized experiment
- assumption: ① no difference between 2 treatments ② ...
- randomization p-value

p-value: prob of getting a test statistic as extreme or more extreme than the observed value of test statistics t^*

$P(T \le t^* | H_0) = \sum \frac{I(t_i \le t^*)}{\binom{N}{N_A}}$ (one sided)

- two sided (not simply double prob of one tail b/c of non-symmetry)

$P(|T - \bar{T}| \ge |t^* - \bar{T}| \,|H_0) = \sum \frac{I(|t_i - \bar{T}| \ge |t^* - \bar{T}|)}{\binom{N}{N_A}}$

- use median to replace mean works as well (also for one/two sided)
- why? mean is affected by extreme data value easily, but median better reflects the "middle class" data feature
- Monte Carlo sampling (when $\binom{30}{\square}$ → large), (randomization practical unless size is small)

one side p-value: $\hat{p} = \frac{1 + \sum_{i=1}^{M} I(t_i \ge t^*)}{M+1}$ including the observed t^* there are $M+1$ test statistics.

- decision Theory

	H_0 true	H_1 true
Accept H_0	correct decision	type II error
Reject H_0	type I error	correct decision

$P(\text{type I}) = \alpha$ $\quad \beta(\text{type II}) = \beta$ $\quad$ power: $1-\beta$ prob rejecting H_0 when H_1 is true.

- exact or conservative, test is guaranteed to control the prob of type I error under very minimal condions: rdmztns of experimental units to the treatments.
- two-sample t-test if $\mu_A, \mu_B \sim N(\mu, \sigma^2)$ with same σ^2

$\bar{Y}_A - \bar{Y}_B \sim N(\mu_A - \mu_B, \sigma^2(\frac{1}{n_A} + \frac{1}{n_B}))$, $\frac{\bar{Y}_A - \bar{Y}_B - \delta}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \sim N(0,1)$, $\delta = \mu_A - \mu_B$

$\frac{\bar{Y}_A - \bar{Y}_B - \delta}{s\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \sim t_{n_A + n_B - 2}$ two-sample t-test $t^* = \frac{\bar{Y}_A - \bar{Y}_B}{s\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}$

What about qq plot? (if data points along qq line ⇒ normality is satisfied!)

- randomization & 2-sample t-test have almost identical p-values.
  ⇒ exchangability is true!
- randomized paired design. (flip a coin!)

comparison

$P(D \le d^* | H_0) = \sum_{i=1} \frac{I(d_i \le d^*)}{N}$ $\quad d^*$: observed diff $\quad D: \bar{A} - \bar{B}$

- when asked "is observed difference due to random chance".
- $Z_\alpha = Z_{0.05} = 1.65$
  $Z_{0.05/2} = 1.96$

地址：南京市察哈尔路37号 电话Tel: (025)83469003 传真Fax: (025)83425737
ADD: 37 Chahaer Road Nanjing http: //www.nsfz.net 邮编：210003

南京师范大学附属中学
HIGH SCHOOL AFFILIATED TO NANJING NORMAL UNIVERSITY

- When # of ways to split data is large, use MC.
- Experiment (if assignment mechanism is controlled by researchers)
- If an subject receives both treatment in a paired fashion ⇒ r.p. design.

For clinical trial: we want to find if there is causation.

- randomization: ensure that the groups will be similar w.r.t. all factors measured in the study & all factors that are not measured.

Power and Sample size.

- one sample z-test: reject at α iff $\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$ (two-sided)

$1-\beta = P(\text{Reject } H_0 \mid \mu=\mu_1) = P(|\bar{X}-\mu_0| \geq \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \mid \mu=\mu_1) = P(\bar{X}-\mu_0 \geq \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \mid \mu=\mu_1) + P(\bar{X}-\mu_0 < \frac{-\sigma}{\sqrt{n}} z_{\alpha/2} \mid \mu=\mu_1)$

$= 1-\Phi\left(z_{\frac{\alpha}{2}} - \left(\frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}\right)\right) + \Phi\left(-z_{\frac{\alpha}{2}} - \left(\frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}\right)\right)$,

power: prob of z-test will detect a mean of μ_1 when $n=$, $\sigma=$, $\alpha=$...

- one sample t-test: reject at α iff $\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right| \geq t_{n-1,\alpha/2}$, similarly for $1-\beta=\cdots$
- two sample t-test: $\bar{Y}_k = \frac{1}{n_k}\sum Y_{ik}$ sample mean. $S_p^2 = \frac{1}{n_1+n_2-2}\sum_{k=1}^{2}\sum_{j=1}^{n_k}(Y_{ik}-\bar{Y}_k)^2$ $\theta = \mu_1-\mu_2$

2-sample t-statistic $T_n = \frac{\bar{Y}_1-\bar{Y}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$

$T_n \sim \underbrace{t_{n_1+n_2-2}}_{df}$ under H_0 & $t_{n_1+n_2-2,\gamma}$ with noncentrality parameter $\gamma = \frac{\theta}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ under H_1

$1-\beta = P(t_{n_1+n_2-2,\gamma} \geq t_{n_1+n_2-2,\frac{\alpha}{2}}) + P(t_{n_1+n_2-2,\gamma} < -t_{n_1+n_2-2,\alpha/2})$

- power as a function of sample size & sd.

ES: $\frac{\mu_1-\mu_2}{\sigma}$, SML: 0.2, 0.5, 0.8 type of allocations: { randomized / nonrandomized

- sample size (known variance & equal allocation)

two-sample z-test .... (similar to two-sample t-test)

solve $z_\beta + z_{\alpha/2} = \left(\frac{|\theta_1|}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right)$ assume $n_1=n_2=\frac{n}{2} \Rightarrow n = \frac{4\sigma^2(z_\beta+z_{\frac{\alpha}{2}})^2}{\theta^2}$

unequal allocation: $r = n_1/n_2 \Rightarrow n_1 = r\cdot n_2 \Rightarrow n_2 = \frac{(1+\frac{1}{r})\sigma^2(z_\beta+z_{\frac{\alpha}{2}})^2}{\theta^2}$

imbalance ⇒ loss of power.

- Binary with proportion: $n_2 = \frac{(z_{\alpha/2}+z_\beta)^2}{\theta^2}\left(p_1(1-p_1)\frac{1}{r} + p_2(1-p_2)\right)$
- power by simulation:
- two sample test assumption: one is iid $N(\mu_1,\sigma^2)$, one is iid $N(\mu_2,\sigma^2)$
- To increase power? either increase n, α or θ or decrease σ.

recommend $n\uparrow$ b/c α, θ are set (stat. meaningful)

$\sigma\downarrow$ is not practical.

地址：南京市察哈尔路37号
ADD: 37 Chahaer Road Nanjing
电话Tel: (025)83469003
http: //www.nsfz.net
传真Fax: (025)83425737
邮编：210003

南京师范大学附属中学
HIGH SCHOOL AFFILIATED TO NANJING NORMAL UNIVERSITY

- causal inference
could have · zero causal effect but + predictive comparison
· positive causal effect but 0 predictive comparison
- fundamental problem of causal inference. *
1. The definition of the causal ~~inference~~ effect depends on the potential outcomes, but not depends on which outcome is actually observed.
2. The causal effect ~~of~~ is the comparison of potential outcomes, for the same unit at the same moment in time post-treatment.
∴ The fundamental problem of CI is at most one of 2 potential outcomes can be observed! Can NEVER measure a causal effect directly.
Then HOW? ① close substitutes / ② Randomization & experiment / ③ statistical adjustment
② do average treatment effect $\bar{y}^1 - \bar{y}^0$
- Stable Unit Treatment Value Assumption (≡ SUTVA): The potential outcomes for any unit do not vary with treatments assigned to other units, & for each unit, there are no different forms or versions of each treatment level, which lead to different outcomes.
- observational studies (not assigned, e.g. lung cancer vs. smoking), sometimes inside an experiment
- If treatment assignment T is conditionally indpdent of ~~=~~ y^0, y^1 given confounding covariates X ⇒ treatment assignment is ignorable. ~~then~~ $y^0, y^1 \perp T | X$
~~take "teaching"~~
non-ignorable: e.g. teaching experience, teacher's motivation ...
- If treatment assignments depend on info not included in the model then watch out!

- propensity score. $e(\vec{x}) = P(T=1|\vec{x})$, $\vec{x}$ are obs. covariates. (before treatment)
↳ known in experiments
↳ unknown in obs. studies ⇒ estimated by models ~~=~~ e.g. logistic regression.
- $\log\left(\frac{p_i}{1-p_i}\right) = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3}$ where $p_i = P(T_i = 1)$
$$\hat{p}_i = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \cdots)}{1 + \exp(\cdots)}$$
- balancing property of propensity score: $P(\vec{x}|T=1, e(\vec{x})) = P(\vec{x}|T=0, e(\vec{x}))$ or $T \perp \vec{x} | e(\vec{x})$
- ignorable treatment assignment $P(T|Y(0), Y(1), \vec{x}) = P(T|\vec{x})$

地址：南京市察哈尔路37号
ADD: 37 Chahaer Road Nanjing
电话Tel: (025)83469003
http: //www.nsfz.net
传真Fax: (025)83425737
邮编：210003

南京师范大学附属中学
HIGH SCHOOL AFFILIATED TO NANJING NORMAL UNIVERSITY

ANOVA (analysis of variances)
- compare multiple samples.
- between treatment variation & within treatment variation

$y_{ij} - \bar{y}_{..} = (y_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})$ y_{ij} ... jth obs under treatment $i = 1, \cdots, a$

$E(y_{ij}) = \mu_i = \mu + \tau_i$, $Var(y_{ij}) = \sigma^2$. τ_i is ith treatment effect.

$y_{i.} = \sum_{j=1}^{n} y_{ij}$, $\bar{y}_{i.} = y_{i.}/n$

$y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$, $\bar{y}_{..} = y_{..}/N$ $\quad N = a \cdot n$

$H_0: \mu_1 = \cdots = \mu_a$ (treatment)
$H_a: \mu_i \neq \mu_j$, $i \neq j$

- ANOVA identity

total sum of squares: $SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = n\sum_{i=1}^{a}(y_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2$

$= SS_{Treat} + SS_{Error}$

$df: N-1$ $\quad df: a-1$ (a - treatment lvls) $\quad df = a(n-1) = an - a = N - a$ (within each trtmt: $n-1$ df, times a $= a(n-1)$)

$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2$

sample variance of ith treatment $s_i^2 = \frac{\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2}{n-1}$

sum up all s_i^2 -> population var $\frac{SS_E}{N-a}$ so SS_E is pooled est. of common var σ^2 within each a treatments

mean square for treatment $MS_{Treat} = \frac{SS_{Treat}}{a-1}$, mean square for error: $MS_E = \frac{SS_E}{N-a}$

$\frac{SS_{Treat}}{\sigma^2} \sim \chi^2_{a-1}$, $\frac{SS_E}{\sigma^2} \sim \chi^2_{N-a}$, thus for H_0: $F = \frac{MS_{Treat}}{MS_E} \sim F_{a-1, N-a}$. (*)

- General ANOVA

Soure of variation	Df	SS	MS	F
B/w treatments	$a-1$	SS_{Treat}	MS_{Treat}	$F = \frac{MS_{Treat}}{MS_E}$
Within treatments	$N-a$	SS_E	MS_E	

- Assumptions: ① additive model $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ (additive model)
(behind p-values in ANOVA) ↳ treatment effect $\tau_i = \mu_i - \mu$

② errors ε_{ij} iid with common var $Var(\varepsilon_{ij}) = \sigma^2$ $\forall i, j$. (constant var)

$E(MS_{Treat}) = \sum_{i=1}^{a}\tau_i^2 + \sigma^2$, $E(MS_E) = \sigma^2$

If no difference between treats means $\Rightarrow \tau_1 = \cdots = \tau_4$, both MS_E, MS_T estimate σ^2

③ if $\varepsilon_{ij} \sim N(0, \sigma^2)$ then MS_T & MS_E independent (normality)

Under H_0 that $\sum_{i=1}^{a}\tau_i^2 = 0$, $F = \frac{MS_T}{MS_E}$ is ratio of 2 indpdt est of σ^2 => (*)

- Least square est. $y_{ij} = \mu + \tau_1 X_1 + \tau_2 X_2 + \tau_3 X_3 + \varepsilon_{ij}$, $X_i = \begin{cases} 1 & \text{if diet B} \\ 0 & \text{o.w.} \end{cases}$ ..., C, D.

so lse: $\hat{\mu} = \bar{y}_{1.}$, $\hat{\tau}_1 = \bar{y}_{2.} - \bar{y}_{1.}$, $\hat{\tau}_2 = \bar{y}_{3.} - \bar{y}_{1.}$, $\hat{\tau}_3 = \bar{y}_{4.} - \bar{y}_{1.}$.

- But NO ASSUMPTIONS behind calculations for SS, df or MS.

地址：南京市察哈尔路37号
ADD: 37 Chahaer Road Nanjing
电话Tel: (025)83469003
http: //www.nsfz.net
传真Fax: (025)83425737
邮编：210003

PRACFINAL

Date

~~2.~~

~~(a) $\frac{1}{2}$~~

~~(b) $\frac{\lambda_1 f_1(x_1,x_2)}{\lambda_1 f_1(x_1,x_2)+\lambda_2 f_2(x_1,x_2)} = \frac{\frac{1}{3}\cdot\frac{1}{\pi}}{\frac{1}{3}\cdot\frac{1}{\pi}+\frac{2}{3}\frac{2(x_1^2+x_2^2)}{\pi}} = \frac{1}{1+4(x_1^2+x_2^2)}$~~

~~$0<x_1^2+x_2^2<1 \Rightarrow 1 > P > \frac{1}{5}$~~

~~$\frac{1}{5}<P<1$~~

Prac final.

Q1.

(a) $\hat{R} = LL^T + \Psi$

$$= \begin{pmatrix} 0.628 \\ 0.786 \\ 0.781 \\ 0.437 \\ 0.404 \end{pmatrix} \begin{pmatrix} 0.628 & \cdots \rightarrow \end{pmatrix} + \begin{pmatrix} 0.605 & & & & 0 \\ & 0.383 & & & \\ & & 0.389 & & \\ & & & 0.084 & \\ 0 & & & & 0.089 \end{pmatrix}$$

$$= \begin{pmatrix} 5 \times 5 \end{pmatrix}$$

appropriate since p-value $= 0.761 > 0.05$

(b) $QQ^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$ ✓

orthogonal.

$L \rightarrow LQ$.

$L = \begin{pmatrix} \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \end{pmatrix}$ $LQ = ?$ ✓

ca2.

Q2. $p(\text{misclassification}) = \frac{p(f_1, x_1>0)}{p(x_1>0)} = \frac{\frac{1}{3}\cdot\frac{1}{2}}{\frac{1}{2}} = \frac{1}{3}$

~~$= p(\text{treat } x \text{ from } f_2 \mid x \text{ from } f_1, x_1>0)$~~

~~$= p(\text{believe it from } f_2 \mid \text{actually from } f_1, x_1>0)$~~

~~$= p(x_1>0 \mid \text{actually from } f_1, x_1>0)$~~

$=$

$i=1$
$j=2$

(b) Classify $\underline{x}$ to group j if $\forall i \neq j$.

$$\lambda_i f_i(\underline{x}) < \lambda_j f_j(\underline{x})$$

$$\frac{f_j(\underline{x})}{f_i(\underline{x})} > \frac{\lambda_i}{\lambda_j}$$

$$\frac{f_2(\underline{x})}{f_1(\underline{x})} = 2(x_1^2+x_2^2) > \frac{\lambda_1}{\lambda_2} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$$

$\Rightarrow x_1^2+x_2^2 > \frac{1}{4} \Rightarrow$ Group 2.

o.w. Group 1.

(c). Nothing b/c no matter what signs X_1 & X_2 are, we are only interested in X_1^2 & X_2^2, which must be non-negative.

So it's useless here.

Q3. PCA vs ICA

same goal different results.

PCA ↓ covariance of the data, ICA ↓ higher order stats e.g. Kurtosis. thus ↓ the mutual info of the output, (what's left is "independent").

PCA outputs orthogonal vectors which are responsible for large portion of variances of the data. ICA identifies indpdt components for non-normal data.

Date

Besides, ICA has 2 ambiguities:
so ICA model equation is underdetermined system,
one cannot determine variances of the indpdt components.
Also. cannot rank the order of dominant component.

Q4. How PCA can be used to construct est. for factor analysis.

Q5.

(a). 2&3 in 1st clustering ✓

(b). Single linkage

	1	2	3	4
2	5			
3	9	(3)		
4	10	7	4	
5	16	11	8	6

	1	23	4
23	5		
4	10	(4)	
5	16	8	6

	1	234
234	(5)	
5	16	6

1234 vs 5

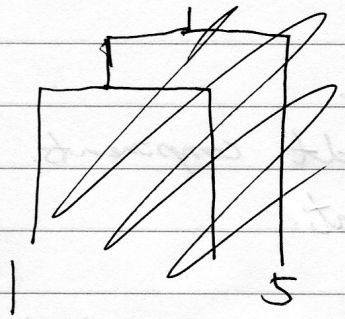
complete linkage

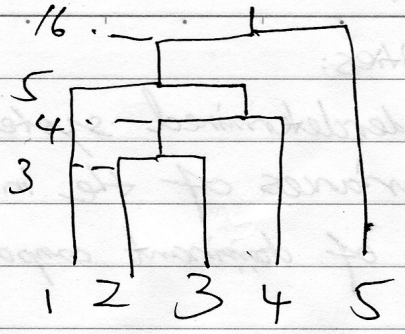
	1	2	3	4
2	5			
3	9	3		
4	10	7	4	
5	(16)	11	8	6

	15	2	3	4
2	(11)			
3	9	3		
4	10	7	4	

	128	3	4
3	(11)		
4	10	4	

1253 vs 4





Single.

Similarly

STA 305 Final Web Notes Review April 12th, 2016

Week 7

- Qualitative Predictors in Regression models.
  - ~~Dummy coding~~ | Dummy coding | (-1, 1, 0).
  - k independent categories then k-1 dummy variables are needed.

e.g. $X = \begin{cases} 1 & \text{if treatment A} \\ 0 & \text{o.w.} \end{cases}$

- dummy coding compares each level to the reference level.
- The intercept is the mean of the reference group.
- Deviation coding
  - coding system compares the mean of the dependent variable for a given level to the overall mean of the dependent variable.

e.g. $X_1 = \begin{cases} 1 & \text{green} \\ -1 & \text{yellow} \\ 0 & \text{o.w.} \end{cases}$ $X_2 = \begin{cases} 1 & \text{pink} \\ -1 & \text{yellow} \\ 0 & \text{o.w.} \end{cases}$ $X_3 = \begin{cases} 1 & \text{purple} \\ -1 & \text{yellow} \\ 0 & \text{o.w.} \end{cases}$

- Estimating treatment effects using least squares.

y_{ij} : ith trt, jth obs.

$$y_{ij} = \mu + \tau_1 X_{i1} + \tau_2 X_{i2} + \tau_3 X_{i3} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

$E(y_{Aj}) = \mu_A = \mu$
$E(y_{Bj}) = \mu_B = \mu_A + \tau_1 \Rightarrow \tau_1 = \mu_B - \mu_A$
$E(y_{Cj}) \Rightarrow \tau_2 = \mu_C - \mu_A$
$E(y_{Dj}) \Rightarrow \tau_3 = \mu_D - \mu_A$

least square est. $\hat{\mu} = \bar{y}_{1\cdot}$, $\hat{\tau}_1 = \bar{y}_{2\cdot} - \bar{y}_{1\cdot}$, $\hat{\tau}_2 =$ ..., $\hat{\tau}_3 =$ ...

- Estimating parameters using least squares (deviation coding):

$E(y_{Aj}) = \mu_A = \tau_0 + \tau_1$
$E(y_{Bj}) = \mu_B = \tau_0 + \tau_2$
$E(y_{Cj}) = \mu_C = \tau_0 + \tau_3$
$E(y_{Dj}) = \mu_D = \tau_0 - \tau_1 - \tau_2 - \tau_3$

$\Rightarrow$ $\tau_0 = ?$, $\tau_1 =$, $\tau_2 =$, $\tau_3 =$

- Multiple comparisons

Suppose k-3 separate (independent) hypothesis level α tests conducted

H_0 true, $P(\text{reject } H_0) = \alpha = 1 - P(\text{do not reject } H_0) = 1 - \alpha$

H_0 is true, $P(\text{reject at least one } H_{0k}) = 1 - (1-\alpha)^3$

family-wise error rate.

But pair-wise error rate is $P(\text{reject } H_{0k}) = \alpha$ for any c.

- When groups sign. diff from ANOVA:
  - testing all pairs ↑ type I error rate
  - beyond (0.25) that a significant difference is detected when the truth is that no difference exists.

Bonferroni Method

two sample t-test. $t_{ij} = \frac{\bar{Y}_{j\cdot} - \bar{Y}_{i\cdot}}{\hat{\sigma}\sqrt{\frac{1}{n_j} + \frac{1}{n_i}}}$ $\quad \hat{\sigma} = \sqrt{MS_E}$ from ANOVA table

- trt i & j sig. diff if $|t_{ij}| > t_{N-k, \alpha/2}$
- A $100(1-\alpha)\%$ simultaneous CI for c pairs $\mu_i - \mu_j$ is

$$\bar{y}_{j\cdot} - \bar{y}_{i\cdot} \pm t_{N-k, \alpha/2c}\hat{\sigma}\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

Tukey Method

- The only diff (between T & B) is choice of critical value.

sig diff if $|t_{ij}| > \frac{1}{\sqrt{2}} q_{\alpha, k, N-k}$

↓

upper α percentile of the Studentized range distn with parameters k & $N-k$ degrees of freedom

CI: $\bar{y}_{j\cdot} - \bar{y}_{i\cdot} \pm \frac{1}{\sqrt{2}} q_{k, N-k, \alpha}\hat{\sigma}\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$

- B is more conservative than T, so CI of T is shorter.

Sample size for ANOVA

test rejects at α-level if $\frac{MS_{Treat}}{MS_E} \geq F_{k-1, N-k, \alpha}$

power $1 - \beta = P(MS_{Treat}/MS_E \geq F_{k-1, N-k, \alpha})$

- Interpret parameters in the additive model for ANOVA $Y_{ti} = \mu + \tau_t + \varepsilon_{ti}$

↓

deviation produced by trt t.

$\hat{\mu} = \bar{y}_1$. $\hat{\mu} + \hat{\tau}_t = \bar{y}_t$, $t = 1, 2, 3, 4$

$\bar{y}_t = \sum_{k=1}^{n_t} y_{tk}/n_t$

∴ least square est. $\hat{\tau}_t = \bar{y}_t - \bar{y}_1$, $t = 1, 2, 3, 4$ $\quad \varepsilon_{ti} \sim N(0, \sigma^2)$

linear model for factorial design:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1} x_{i2} + \cdots + \varepsilon_i$$

Week 8

factorial design

- fixed # of levels of each of a number of factors
  - 2^3: 2 levels, 3 factors - 8 experimental conditions
- can involve ~~dif~~ factors ~~of~~ with different # of levels.
- ANOVA compares the individual experimental conditions with each other
- factorial --- generally compares combinations of experimental conditions.
- ~~main effect~~
- cube plots $= \bar{y}_+ - \bar{y}_-$
- factorial effects { main effects (average effect of a var)
  interaction effects

- replication in factorial design
  - not always feasible actually
  - variance of each measurement is σ^2.
  - estimated variance at each set of condition is

$$s_i^2 = \frac{(y_{i1} - y_{i2})}{2}$$

Pooled est of σ^2 is $s^2 = \frac{\sum_{i=1}^{8} s_i^2}{8} = \frac{\sum diff^2}{\#}$.

- Estimate of the error variance & standard error of effects from replicated runs

$$Var(\text{effect}) = \left(\frac{1}{8} + \frac{1}{8}\right) s^2$$

$$se = \sqrt{Var}$$

- Interpretation:
  interaction plot (parallel = no interaction)
- Linear model for factorial design
  ☆: The estimated least square coefficients are one-half the factorial estimates, the ~~sample~~ intercept β_0 is the sample mean.
- Advantages of factorial design over one-at-a-time designs
  - if effect is the same, factorial more efficient, requires fewer obs to achieve the same precision.
  - if different, factorial can detect & estimate interactions.

Week 9

- Randomized Block design
  - Comparison of 2 trt:
    - unblocked arrangements: unpaired comparison of 2 trt groups.
    - blocked arrangements: paired comparison of 2 trts
  - Comparison of more trts
    - unblocked : randomized one-way design
    - blocked : randomized block design

- ANOVA for RB design:

~~Say k=1, ANS~~

ANOVA identity.

- a is the number of trts & b # of blocks
- $y_{i\cdot}$: total of all obs. taken under treatment i.
- $y_{\cdot j}$: total of all obs. taken under block j.
- $y_{\cdot\cdot}$: grand total of all obs.
- $N = ab$

$\bar{y}_{i\cdot} = y_{i\cdot}/b$

$\bar{y}_{\cdot j} = y_{\cdot j}/a$

$\bar{y}_{\cdot\cdot} = y_{\cdot\cdot}/N$

$SS_T = SS_{Treat} + SS_{Block} + SS_E$

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

- N obs. SS_T has $N-1$ dof
- a trts & b blocks so SS_{Treat} has $a-1$ & SS_{Block} has $b-1$ dof
- SS_E has $(a-1)(b-1)$ dof

- linear model for RBD: $y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$

  $E(\varepsilon_{ij}) = 0$

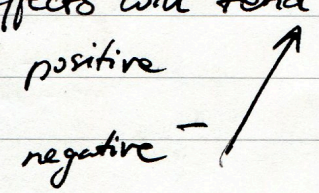
- Latin Square design: { the design allows for blocking with 2 vars; the 2 blocking vars must have the same # of levels

- LSD assumes additive, so no interaction effects.
- Graeco-Latin sq. design.
  (3 blocking vars)  $\alpha, \beta, \gamma, A, B, C \ldots$
- hyper ---- (> 3 blocking vars)
- a balanced incomplete block design has the property that every pair of trts occurs together in a block the same # of times.

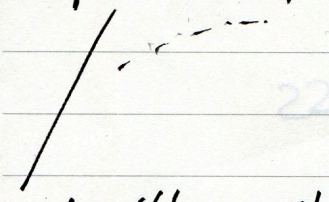
Week 10

- Normal Q-Q plot
  - A marked (systematic) deviation of the pts from the straight line would indicate that:
    - no normality
    - variance is not ~~cost~~ constant

- Normal plot
- When some of the effects are nonzero the corresponding estimated effects will tend to be larger & fall ~~out~~ off the straight line.
- positive +
  negative −



- Half-normal plots



- Lenth's method: testing sig. for experiments without var est.
  $|ME| < |SME|$, SME 나중에 sig
- Blocking factorial designs.
- A general approach for arranging 2^k design in 2^q blocks of size 2^{k-q} is:
  factorial effects V_i, block vars B_q
  $B_1 = v_1, \ldots, B_q = v_q$
  $B_1 B_3 = 245$
  $B_1 B_2 B_3 = 34$
  e.g. $B_1 = 135$, $B_2 = 235$, $B_3 = 1234$, $B_1 B_2 = 12$, $B_2 B_3 = 145$

1. Lower order effects are more likely to be important than higher-order effects.
2. Effects of the same order are equally likely to be important.

- ? Less 2-way better.

- Fractional factorial designs.
  A quarter fraction of the full 2^5 design
  or
  A 2^{5-2} design.
- Main effect E is aliased with BCD interaction.
  E=BCD or I=BCDE
  - factorial effects are missing for effects that are aliased.

Week 11

- Splitplot design is like a RBD.
- 2 sources of variances, σ_w^2, σ_s^2
  ↓ wholeplot (large) ↓ subplot. (small)

Whole plot factor ~~plot~~ A with I levels &
Subplot factor B with J levels
n times replication.

ANOVA.

Source	DF	SS
Wholeplot		
replication	$n-1$	SS_{Rep}
A	$I-1$	SS_A
replication × A (whole plot error)	$(n-1)(I-1)$	SS_W
Subplot		
B	$J-1$	SS_B
A×B	$(J-1)(I-1)$	$SS_{A\times B}$
Subplot error	$I(J-1)(n-1)$	SS_S

linear model of split plot:

$$Y_{ijk} = \eta + \tau_k + \alpha_i + (\tau\alpha)_{ki} + \beta_j + (\alpha\beta)_{ij} + (\tau\beta)_{kj} + (\tau\alpha\beta)_{kij} + \varepsilon'_{ijk}, \quad \varepsilon'_{ijk} \sim N(0, \sigma^2)$$

$i = 1, \cdots, I$,
$j = 1, \cdots, J$
$k = 1, \cdots, n$

τ_k : effect of kth replication
α_i : ith main effect of A
$(\tau\alpha)_{ki}$: (k, i)th interaction effect b/w replication & A
↓ this is whole plot error

β_j : jth main effect of B
$(\alpha\beta)_{ij}$: (i, j)th interaction between A & B
$(\tau\beta)_{kj}$ (k, j)th
$(\tau\alpha\beta)_{kij}$: - -
ε'_{ijk} : error term

Sum = ε_{kij}
↓
subplot error

$\varepsilon_{kij} < (\tau\alpha)_{kj}$ b/c subplot more homogeneous than whole plots.

- Some from Reviews
  - Lenth's plot. use ME, not ~~AE~~SME unless asked to adjust for multiple comparison.

- If the design is replicated then it's possible to calc standard errors of factorial estimates.
  So we can use sig. testing & CI to evaluate the sig. of effects.
  So a graphical aid to assess significance is not required.

- If the interaction is sig. then the main effect should be interpreted in conjunction with the interaction.