- $\boxed{\text{Monte Carlo Sampling:}}$ one-side $p=\frac{1+\sum_{i=1}^{M} I(t_i \geq t^*)}{M+1}$
- decision theory

	H_0 true	H_1 true
Accept H_0	✓	type II error
Reject H_0	type I error	✓

$1-\beta$: power, prob of rejecting H_0 when H_1 true.

- $\boxed{\text{2-sample t-test}}$ if $\mu_A, \mu_B \sim N(\mu, \sigma^2)$

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

$\frac{\bar{Y}_A-\bar{Y}_B-\delta}{S\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}} \sim t_{n_A+n_B-2}$, $t^*=\frac{\bar{Y}_A-\bar{Y}_B}{S\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}}$

randomization & 2-sample t-test almost identical p's $\Rightarrow$ exchangability is true!

randomized $\boxed{\text{paired design}}$: $P(D \leq d^* | H_0) = \sum \frac{I(d_i \leq d^*)}{N}$

$D: \bar{A}-\bar{B}$, d^*: observed diff.

if a subject receives both treatments in a paired fashion..

- clinical trial: random ensures groups will be similar w.r.t all factors measured in the study & all factors not...
- 1-sample z-test (reject H_0 iff $|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}| \geq Z_{\alpha/2}$)

$\boxed{\text{power}}$ $1-\beta = P(|\bar{X}-\mu_0| \geq \sigma/\sqrt{n}\, Z_{\alpha/2} | \mu=\mu_1) = P(\quad \geq \quad)+P(\quad < - \quad)$

$=1-\Phi(Z_{\frac{\alpha}{2}}-(\frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}))+\Phi(-Z_{\frac{\alpha}{2}}-(\frac{\mu_1-\mu_0}{\sigma/\sqrt{n}}))$ prob of test ~~detc~~ detecting a mean of μ_1 when $n, \sigma, \alpha = \ldots$

- $\boxed{\text{2-sample t-test}}$ $\bar{Y}_k=\frac{1}{n_k}\sum Y_{ik}$, $S_p^2=\frac{1}{n_1+n_2-2}\sum_{k=1}^{2}\sum_{i=1}^{n_k}(Y_{ik}-\bar{Y}_k)^2$

$T_n=\frac{\bar{Y}_1-\bar{Y}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$, noncentrality para $\gamma=\frac{\theta}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$, $\theta=\mu_1-\mu_2$

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

- $ES=\frac{\mu_1-\mu_2}{\sigma}$, SML, 0.2, 0.5, 0.8
- sample size (σ^2 & allocation) (2-sample z-test)

$n_1=n_2=\frac{n}{2} \Rightarrow n=\frac{4\sigma^2(Z_\beta+Z_{\frac{\alpha}{2}})^2}{\theta^2}$; $r=\frac{n_1}{n_2} \Rightarrow n_2=\frac{(1+\frac{1}{r})\sigma^2(Z_\beta+Z_{\frac{\alpha}{2}})^2}{\theta^2}$

imbalance $\Rightarrow$ loss of power

- binary with proportion $n_2=\frac{(Z_{\alpha/2}+Z_\beta)^2}{\theta^2}(p_1(1-p_1)\frac{1}{r}+p_2(1-p_2))$
- 2-sample test assumption: one ~ iid $N(\mu_1,\sigma^2)$, one ~ iid $N(\mu_2,\sigma^2)$
- $\uparrow$ power? $n\uparrow$ b/c α, θ set, $\sigma\downarrow$ not practical.
- $\boxed{\text{causal inference}}$: fundamental problem: at most one of ~~treat~~ potential outcomes can be observed. can NEVER measure CI directly. Then ① close substitutes ② randomization & experiment ③ statistical adjustment.
- average treatment effect: $\bar{Y}^1-\bar{Y}^0$
- SUTVA: potential outcomes for any unit do not vary with treatments assigned to other units & for each unit there are no different forms of each treatment level, which lead to different outcomes.
- $\boxed{\text{propensity score}}$: $e(\vec{x})=P(T=1|\vec{x})$. $\vec{x}$ are obs. covariates (before treatment) p.s. is known in exp but unknown in obs. studies $\Rightarrow$ est. by models e.g. logistic regression.

$\log(\frac{p_i}{1-p_i})=\hat{\beta}_0+\hat{\beta}_1 x_{i1}+\cdots+\hat{\beta}_3 x_{i3}$ where $p_i=P(T_i=1)$

- $\boxed{\text{balancing property of p.s.}}$ $P(\vec{X}|T=1, e(\vec{x}))=P(\vec{X}|T=0, e(\vec{x}))$ $\Rightarrow$ randomized!
- $\boxed{\text{ignorable treatment assignment}}$ $P(T|Y(0), Y(1), \vec{x})=P(T|\vec{x})$

$\boxed{\text{ANOVA}}$ between trtmt variation & within trtmt variation

$y_{ij}-\bar{y}_{..}=(y_{i.}-\bar{y}_{..})+(y_{ij}-\bar{y}_{i.})$ y_{ij} ... jth obs in trtmt $i=1,\ldots,a$

$E(y_{ij})=\mu_i=\mu+\tau_i$, $Var(y_{ij})=\sigma^2$

τ_i is treatment effect, $y_{i.}=\sum_{j=1}^{n} y_{ij}$, $\bar{y}_{i.}=y_{i.}\cdot\frac{1}{n}$

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

ANOVA identity

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

$df=N-1$ $\qquad = SS_{Treat} + SS_E$

$\qquad df=a-1 \qquad df=a(n-1)=N-a$

- sample var ith-treatment

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

sum up population var $\frac{SS_E}{N-a}$ ... pooled est. of σ^2 within each a treatment

$MS_{Treat}=\frac{SS_{Treat}}{a-1}$, $MS_E=\frac{SS_E}{N-a}$

$SS_{Treat}/\sigma^2 \sim \chi^2_{a-1}$, $SS_E/\sigma^2 \sim \chi^2_{N-a}$

$F=MS_{Treat}/MS_E \sim F_{a-1, N-a}$

- $\boxed{\text{Assumptions of ANOVA}}$

① Additive model: $y_{ij}=\mu+\tau_i+\varepsilon_{ij}$.

② Constant variance: ε_{ij} iid with common var $Var(\varepsilon_{ij})=\sigma^2$ $\forall i,j$ by plot. If no diff b/w treats $\Rightarrow \tau_1=\cdots=\tau_4$ both MS estimate σ^2.

③ normality: if $\varepsilon_{ij} \sim N(0,\sigma^2)$ then MS_E & MS_{Treat} independent by qq line & qqplot.

- Least square est. $y_{ij}=\mu+\tau_1 X_1+\cdots+\varepsilon_{ij}$, $X_1=\begin{cases}1 & \text{if diet B}\\ 0 & \text{o.w.}\end{cases}$ ...

So $\hat{\mu}=\bar{y}_{1.}$, $\hat{\tau}_1=\bar{y}_{2.}-\bar{y}_{1.}$, ..., $\hat{\tau}_3=\bar{y}_{4.}-\bar{y}_{1.}$

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

$\boxed{MS_E \text{ is unbiased}}$ $y_{it}=\mu+\varepsilon_{it}+\tau_i$, $\varepsilon_{it} \sim N(0,\sigma^2)$

$\Rightarrow y_{it} \sim N(\mu+\tau_i, \sigma^2)$, $\bar{y}_{i.} \sim N(\mu+\tau_i, \sigma^2/n)$

$\sigma^2=E(y_{it}^2)-(\underbrace{\quad}_{\mu+\tau_i})^2$, $\frac{\sigma^2}{n}=E(\bar{y}_{i.}^2)-(\mu+\tau_i)^2$

$SS_E=\sum_i\sum_t(y_{it}^2-2y_{it}\bar{y}_{i.}+\bar{y}_{i.}^2)$

$=\sum_i(\sum_t y_{it}^2-\underbrace{2\sum_t y_{it}\bar{y}_{i.}}_{2\bar{y}_{i.}n\bar{y}_{i.}}+n\bar{y}_{i.}^2)$

$=\sum_i(\sum_t y_{it}^2-n\bar{y}_{i.}^2)$

$E(SS_E)=\sum_i(\sum_t E(y_{it}^2)-nE(\bar{y}_{i.}^2))$

$=\sum_i(n(\sigma^2+(\mu+\tau_i)^2)-n(\frac{\sigma^2}{n}+(\mu+\tau_i)^2))$

$=\sum_i(n-1)\sigma^2=v(n-1)\sigma^2=(N-v)\sigma^2$

$E(MS_E)=E(\frac{SS_E}{N-v})=\sigma^2$

$\boxed{\text{power of z-test}}$ $n=\frac{4\sigma^2}{\theta^2}(Z_{\alpha/2}+Z_\beta)^2 \Rightarrow \sqrt{\frac{\theta^2 n}{4\sigma^2}}=Z_{\frac{\alpha}{2}}+Z_\beta$

$\frac{|\theta|\sqrt{n}}{2\sigma}-Z_{\alpha/2}=Z_\beta$, $\Phi(\frac{\theta\sqrt{n}}{2\sigma}-Z_{\alpha/2})=\Phi(Z_\beta)=1-\beta$

$1-\Phi(\quad)=1-\beta$ since $\Phi(x)=1-\Phi(-x)$.

$\boxed{SS_{Total}=SS_{Treat}+SS_E}$ $y_{ij}-\bar{y}_{..}=(y_{ij}-\bar{y}_{i.})+(\bar{y}_{i.}-\bar{y}_{..})$

Take square both sides.

$SS_{Total}=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{..})^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}[(y_{ij}-\bar{y}_{i.})+(\bar{y}_{i.}-\bar{y}_{..})]^2$

$=\sum_i\sum_j[(y_{ij}-\bar{y}_{i.})^2+2(y_{ij}-\bar{y}_{i.})(\bar{y}_{i.}-\bar{y}_{..})+(\bar{y}_{i.}-\bar{y}_{..})^2]$

$=SS_{Treatment}+0+SS_E$

$\sum_i\sum_j 2(y_{ij}-\bar{y}_{i.})(\bar{y}_{i.}-\bar{y}_{..})=0$

$\sum_i\sum_j 2(y_{ij}\bar{y}_{i.}-\bar{y}_{i.}^2-\bar{y}_{..}y_{ij}+\bar{y}_{i.}\bar{y}_{..})$

$=\sum_i(2n\bar{y}_{i.}^2-2\bar{y}_{i.}^2-2n\bar{y}_{i.}\bar{y}_{..}+2\bar{y}_{i.}\bar{y}_{..})$

$=0$

diff b/w 2-sample & paired experiments

① 2-sample compares 2 independent samples on diff exp. units.

② paired compares two samples on the same exp. unit.

$\boxed{\text{power}}\uparrow$ ① $n\uparrow$ ② $\uparrow\alpha$ ③ $\downarrow\sigma$ ④ $\uparrow\theta$.

Multiple comparison problem

problem of testing more than one hypothesis simultaneously

$\Rightarrow$ theoretical type I error increase $\uparrow$.

The collection of comparisons $\Rightarrow$ "family". family error rate

= prob (at least one comparison will include an type I error)

- compare 3 methods (propensity score)
  similar results for treatment effect.
  unadjusted (estimated)
- matching: for each unit we have a propen. score
  - randomly select one treated subject.
  - match to a control subject with closest P.S.
  - eliminate both from pool until no matching

Y_{ij}: ith trt, jth obs. $Y_{ij} = \mu + \tau_1 X_{i1} + \tau_2 X_{i2} + \tau_3 X_{i3} + \varepsilon_{ij}$, $\varepsilon_{ij} \sim N(0, \sigma^2)$

$E(Y_{Aj}) = \mu_A = \mu$, $E(Y_{Bj}) = \mu_B = \mu_A + \tau_1 \Rightarrow \tau_1 = \mu_B - \mu_A$, ... $\tau_2 = \mu_C - \mu_A$, $\tau_3 = \mu_D - \mu_A$

$\hat{\mu} = \bar{Y}_{1.}$, $\hat{\tau}_1 = \bar{Y}_{2.} - \bar{Y}_{1.}$, ...

(Deviation coding) Estimating parameters $E(Y_{Aj}) = \mu_A = \tau_0 + \tau_1$, $E(Y_{Bj}) = \mu_B = \tau_0 + \tau_2$

... $E(Y_{Dj}) = \mu_D = \tau_0 - \tau_1 - \tau_2 - \tau_3$, $\tau_0 = ?$ ...

Multiple comparison: P(reject H_{0k}) $= \alpha \Rightarrow$ P(reject at least one H_{0k}) $= 1-(1-\alpha)^3$

pairwise error rate is still α

When group significantly diff from ANOVA: testing all pairs ↑ type I error beyond (0.05) that a significant diff is detected when the truth is that no difference exists.

Bonferroni Method: $t_{ij} = (\bar{Y}_{i.} - \bar{Y}_{j.}) / \hat{\sigma}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}$, $\hat{\sigma} = \sqrt{MS_E}$ from ANOVA

sig diff if $|t_{ij}| > t_{N-k, \alpha/2c}$, CI: $\bar{Y}_{i.} - \bar{Y}_{j.} \pm t_{N-k,\alpha/2c}\,\hat{\sigma}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}$

Tukey Method: if $|t_{ij}| > \frac{1}{\sqrt{2}} q_{k,N-k,\alpha}$ → upper α percentile of the studentized range dist'n with k & N-k dof.

B is more conservative than T. CI: $\bar{Y}_{i.} - \bar{Y}_{j.} \pm \frac{1}{\sqrt{2}} q_{k,N-k,\alpha}\,\hat{\sigma}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}$

Sample size for ANOVA. test rejects at α if $MS_{Treat}/MS_E \geq F_{k-1,N-k,\alpha}$ power: $1-\beta = P(MS_{Treat}/MS_E \geq F_{k-1,N-k,\alpha})$

- Factorial design (fixed # of levels of each of a # of factors)
  ANOVA compares the individual with each other
  factorial generally compares the combinations of them.
- replication is not always feasible, replicated runs Var(effect)
- pooled est. of σ^2 $s^2 = \sum diff^2/\#$, $= (\frac{1}{8}+\frac{1}{8})s^2$, $se = \sqrt{Var}$
- The est. coeff are one-half the factorial est. the intercept β_0 is the sample mean.
- interaction plot ..
- Linear model for factorial design
- Advantages of factorial design over one-at-a-time design:
  - if effect is same, factorial more efficient, requires fewer obs to achieve the same precision.
  - if effect different, factorial can detect & estimate interactions.
- Randomized Block design: $\underset{N-1}{SS_{Total}} = \underset{a-1}{SS_{Treat}} + \underset{b-1}{SS_{Block}} + \underset{(a-1)(b-1)}{SSE}$

linear model:

$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$, $E(\varepsilon_{ij}) = 0$.

Latin square: the design allows for blocking with 2 vars
the 2 blocking vars must have the same # of levels.
LSD assumes additive

Graeco-Latin sq design (3 blocking vars) ≥3 .. hyper

A balanced incomplete block design has the property that every pair of trt occurs together in a block the same # of times.

A marked (deviation) in Q-Q plot ⇒ ① no normality ② variance is not constant.

Normal / Half-normal plot / Lenth's plot, |SME| 以上 or |ME| 以上.

Blocking factorial design

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.

A general approach for arranging 2^k design in 2^q blocks of size 2^{k-q} is: factorial effects V_i, block effect B_q ... (don't confound with main effect!)

- Fractional factorial designs: A quarter fraction of full 2^5 design $= 2^{5-2}$
- Fractional factorial design. Main effect can be aliased with interaction effect. factorial effects are missing for e.g. E=BCD, then I=BCDE. effects that are aliased.
- defining relation & aliasing relations

splitplot design is a RBD: 2 sources of variation, σ_w^2 (whole plot), σ_s^2 (subplot)
wholeplot factor A with I levels & subplot ~~with~~ 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 (wholeplot 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 splitplot

$Y_{ijk} = \eta + \tau_k + \alpha_i + (\tau\alpha)_{ki} + \beta_j + (\alpha\beta)_{ij} + (\tau\beta)_{kj} + (\tau\alpha\beta)_{kij} + \varepsilon'_{ijk}$,

$i = 1, \dots, I$, $j = 1, \dots, J$, $k = 1, \dots, n$.

τ_k: effect of kth replication
α_i: ith main effect of A
$(\tau\alpha)_{ik}$: (k, i)th interaction effect b/w replication & A (whole plot error)
β_j: jth main effect of B
$(\alpha\beta)_{ij}$: (i, j)th interaction b/w A & B
$(\tau\beta)_{kj}$: (k, j)th interaction b/w replication & B
$(\tau\alpha\beta)_{kij}$: (k, i, j)th interaction b/w rep, A & B
ε': error term

} ε_{kij} (subplot error) b/c homogeneity.

- Lenth's plot. use ME, not 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
- Design matrix e.g.

	A	B	C	D
1	1	1	1	1
2	1	1	-1	-1
3	1	-1	1	-1
4	-1	1	1	-1
5	-1	-1	1	1
6	-1	1	-1	1
7	1	-1	-1	1
8	-1	-1	-1	-1

D=ABC.

again: factorial effect is 2 times' linear model est.

- (blocks of size 2) paired design since each sample is weighed on 2 scales.
- Write 2^3 design in blocks of size 2 s.t. no main effect is confounded with block differences.
  $B_1 = 12$, $B_2 = 13$ Block I: $B_1 = +$, $B_2 = -$ (2 runs)
  II: $B_1 = +$, $B_2 = -$ (2 runs)
  III: $B_1 = -$, $B_2 = +$ (2 runs)
  IV: $B_1 = -$, $B_2 = -$ (2 runs)
  Block effect: $B_1 \cdot B_2 = 12\,13 = 23$ (not confounded with main)
- 2^4 in 4 blocks: $B_1 = 124$, $B_2 = 134$ ...
  $B_1 \cdot B_2 = 124 \cdot 134 = 23$ (no confounded ...)
- No paired design ⇒ since dependency. A design that is appropriate randomized subjects to one of the 2 trts since → response that are independent.
- split plot design: whole plot: "hard to change" factor
  subplot: ...
- model (with 3 parts) for analysis:
- The main effect of preparation method can be evaluated with an F test, observed F-values is $MS_{prep}/MS_{prep:day} = 7.077$, refer to $F_{2,4}$ distribution. Critical is 6.944. So sig. diff.
  ↓ $F_{3-1,4}$

The main effect of temp is evaluated using subplot error.
ANOVA: Error within, obs. F-value is $36.427 = F_{3,18}^*$
p-value is 0. so sig. difference.

- CI contains 0 (Tukey pairwise) ⟺ ~~difference~~ not significantly different from 0. ...
- table of contrast of a factorial design has k! columns
- sometimes, the significance results from plot (say, interaction plot) & hypothesis testing are not consistent,
- Latin square (linear model)
  $y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \varepsilon_{ijk}$, $\varepsilon_{ijk} \sim N(0, \sigma^2)$
- RBD: ANOVA identity in details:
  $SS_T = SS_{Treat} + SS_{Blocks} + SS_E$
  $$\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{..})^2 = b\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 + a\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$
- For fractional factorial
  How to get aliased relations?
  Use I to multiple every possible main & interaction effects!
- Fractional factorials use the redundancies by arranging that lower order effects are confounded with higher order interactions that are assumed negligible.
- ANOVA: (General)

Source	Dof	SS	MS	F
Between Trts	$a-1$	SS_{Treat}	MS_{Treat}	$F_{obs} = \frac{MS_{Treat}}{MS_E}$
Within Trts	$N-a$	SS_E	MS_E	

↙ compare with $F_{a-1,N-a}$

~~Subplot~~ split plot design consideration:
feasibility & efficiency! Whenever there are factors difficult to change & others easy to change.

ANOVA for split plot has 3 parts:
First one is MS error for one-way ANOVA
Second is wholeplot error.
Third is subplot error.

HW3 ~~第~~ Q4:
$Var(\bar{y}_1 - \bar{y}_2) = 4\sigma^2 \cdot \frac{1}{16} + \frac{4\sigma^2}{16} = \frac{\sigma^2}{2}$
- main effect of A in 2^2 replicated design: $A = (-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 + \bar{y}_4)/2$
$Var(A) = Var((-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 + \bar{y}_4)/2) = \frac{1}{4}(4\sigma^2/2) = \frac{1}{2}\sigma^2$

e.g. of interpretations of main/interaction effects:
- main effect of °C is average difference between yield when °C is at low level vs. yield when °C is high.
- interaction effect (2-way) is the average difference between effect of °C when pH is high vs. ~~ee~~ effect of °C when pH is low,
- 3-way: average difference between interaction ? when ? is low vs. interaction ? when ? is high,

~~*~~. For linear model of factorial design,
main effect calculation. divisor is one-half of total runs: transforming the contrast into a difference between 2 averages.

F second error dof - -

F third error dof $\underset{\text{main}}{-}\ \underset{\text{residual}}{-}$