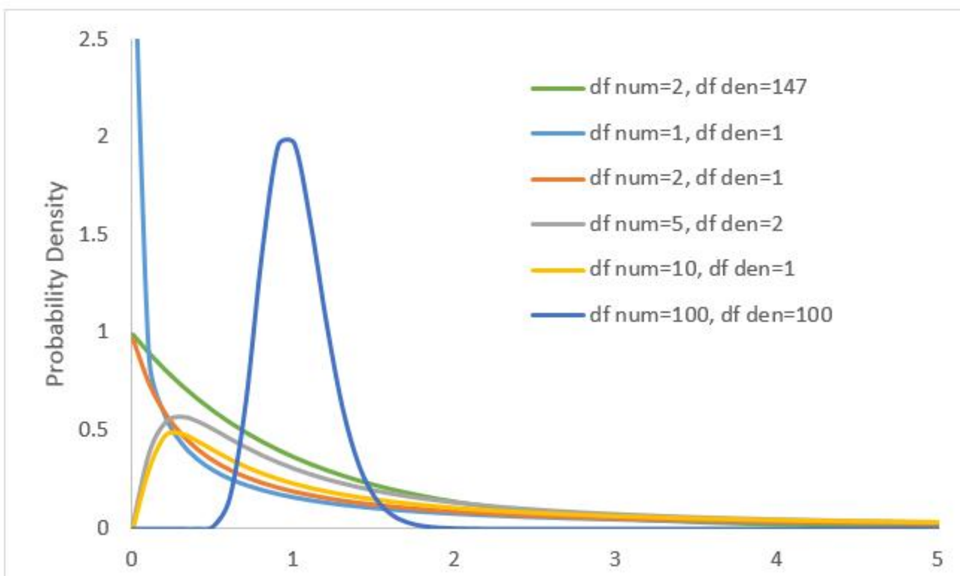
1-way ANOVA : fixed effects model

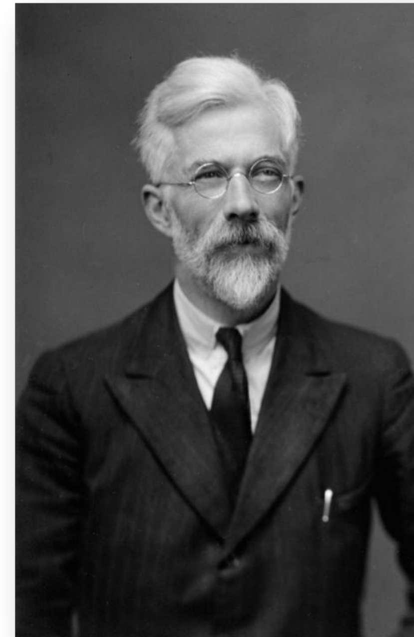
if $X \sim x_n^2, Y\sim x_m^2$. X and Y are independent.

$$F = \frac{X/n}{Y/m}, then$$

$$F \sim F_{n,m}, or F \sim F(n,m)$$

The PDF of an F random variable with r1 numerator degrees of freedom and r2 denominator degrees of freedom is:





Ronald A. Fisher

OVERVIEW

Properties of chi-square distribution

(3) If $X \sim N(\mu, \sigma^2), (X_1, \dots, X_n)$ are randomly selected from population X

The distribution of $Y \sim \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2$

$$because\ X \sim N(\mu, \sigma^2),\ so\ \frac{X_i - \mu}{\sigma} \sim N(0,1)$$

According to the definition of chi-square distribution

$$Y = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim x_n^2$$

Properties of chi-square distribution

(4)

$$x_n^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 \approx \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2}$$

$$S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$$

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = (n-1) * S^2$$

$$x^2 \approx \frac{(n-1) * S^2}{\sigma^2} \sim x_{n-1}^2$$

MOTIVATIONS

Practical demands

- In health sciences, researchers often need to determine **if multiple groups have different means**.
- Formally, they want to test the null hypothesis of no difference among the means, μ_i's, of $k(\geq 3)$ groups:

$$H_0: \mu_1 = \ldots = \mu_k \text{ vs. } H_1: \text{Otherwise}$$

(namely, at least two of the μ_i's are different)

Limitations of *t* tests

Two-sample *t* tests enable us to *judge* if two groups have different means. But these tests *are inefficient and insufficient* to determine if *three or more groups differ in means*.

It is of *clear drawbacks* to perform *all possible separate pairwise comparisons* by the *t* tests.

(1) It requires a *substantial amount of work* to carry out many *t* tests, particularly for large data sets.

(2) A more severe consequence is that *it is very likely to lead to a false conclusion*.

Suppose we draw k independent samples from k populations with an identical mean (e.g., $N(0, i)$ for $i = 1, \ldots, k$). In such a scenario, the joint $H_0: \mu_1 = \ldots = \mu_k$ is true.

The number of all possible pairwise t tests is $C_k^2 = k(k-1)/2$, e.g., $C_5^2 = 10$. Let the nominal significance level $\alpha = 0.05$ for each t test. Assuming independence between pairwise t tests, it can be shown that the probability of rejecting the overall null is $\mathbf{1 - (1 - 0.05)^{C_k^2}}$.

This is the type I error rate of pairwise comparisons procedure. As long as $k > 2$, the procedure cannot control type I error rate at α (**Figure1.1 in next slide**). Testing all possible pairs of means may result in severe type I error rate.

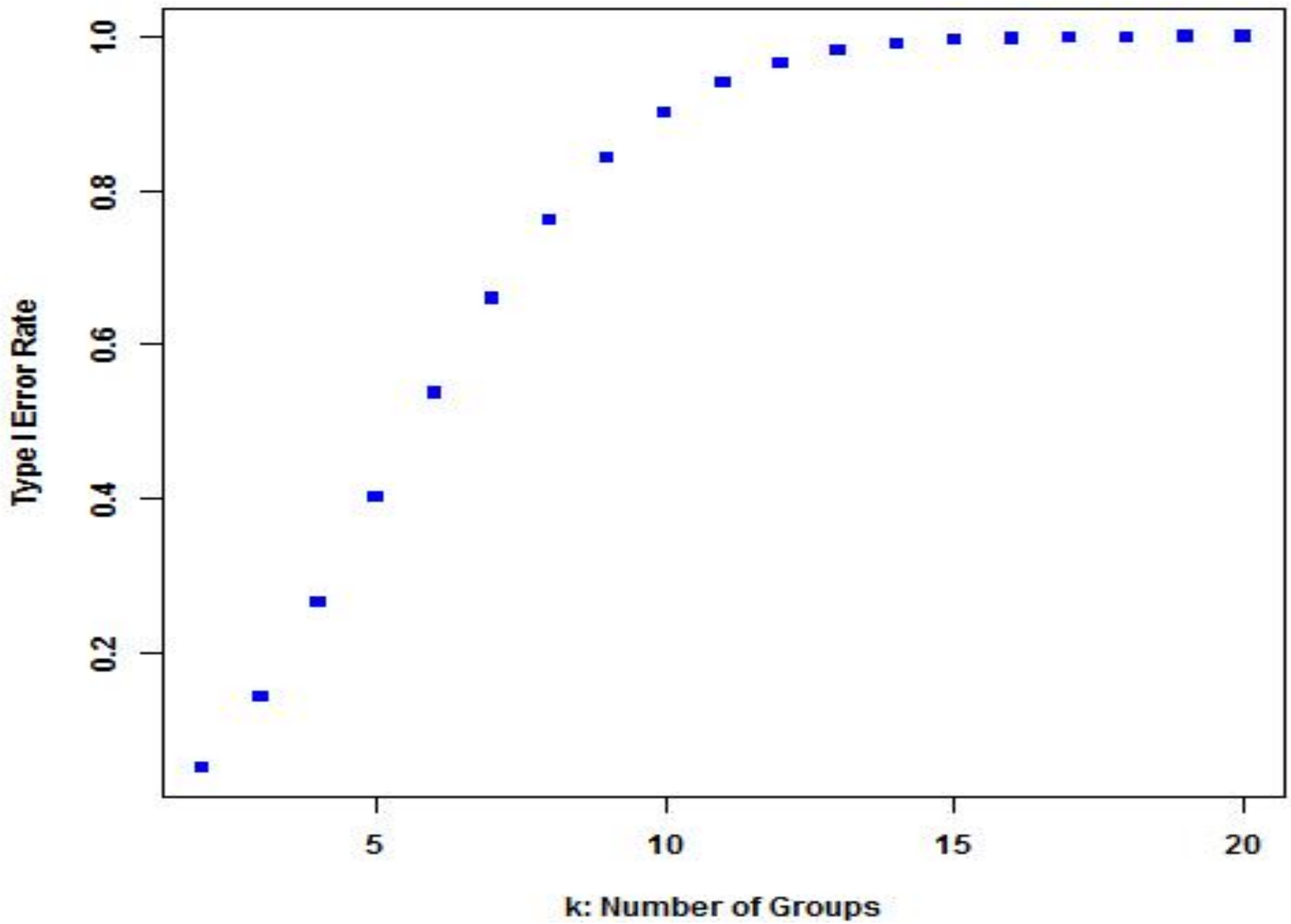


Figure 1.1: Type I error rate of the pairwise comparisons procedure. For each pairwise comparison, the nominal level is $\alpha = 0.05$.

CONTENTS

- Purpose and logic of 1-way ANOVA
- 1-way ANOVA for CRD
- ANOVA for RBD
- Post-hoc analyses

1. PURPOSE AND LOGIC OF 1-WAY ANOVA

(1) Understand in a simpler way

Example 1.1: Nine individuals were randomly assigned to take different drugs, and then the serum levels of a special protein were tested.

Drug A	Drug B	Drug C
3	5	5
2	3	6
1	4	7
$\overline{X1}$=2	$\overline{X2}$=4	$\overline{X3}$=6

The ground (overall) mean: $\overline{\overline{X}}$=(3+2+1+5+3+4+5+6+7)/9=4

$random\ error$

Not every value within the same drug group are identical: i.eg, 3$\neq \overline{X1}$, 2$\neq \overline{X1}$, 1$\neq \overline{X1}$

Why?

$effect\ of\ different\ drugs$

group means are not identical:
$\overline{X1}$, $\overline{X2}$ and $\overline{X3}$ $\neq \overline{\overline{X}}$; $\overline{X1} \neq \overline{X2} \neq \overline{X3}$

Interesting

ith group

jth observation

Drug A	Drug B	Drug C
3	5	5
2	3	6
1	4	7
$\overline{X1}$=2	$\overline{X2}$=4	$\overline{X3}$=6

The ground (overall) mean: $\overline{\overline{X}}$=4

Sum of squares total=$(3-4)^2+(2-4)^2+(1-4)^2+(5-4)^2+(3-4)^2+(4-4)^2+(5-4)^2+(6-4)^2+(7-4)^2$ =30

$$SST=\sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij}-\overline{\overline{X}})^2$$

Df=3×3-1

Sum of squares within= $(3-2)^2+(2-2)^2+(1-2)^2+(5-4)^2+(3-4)^2+(4-4)^2+(5-6)^2+(6-6)^2+(7-6)^2$ =6

$$SSW=\sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij}-\overline{X}_i)^2$$

Df=3×(3-1)

Sum of squares between= $3*(2-4)^2+3*(4-4)^2+3*(6-4)^2$=24

$$SSB=\sum_{i=1}^{k}\sum_{j=1}^{ni}(\overline{X}_i-\overline{\overline{X}})^2=\sum_{j=1}^{ni}n_i(\overline{X}_i-\overline{\overline{X}})^2$$

Df=3-1

SST=SSW+SSB

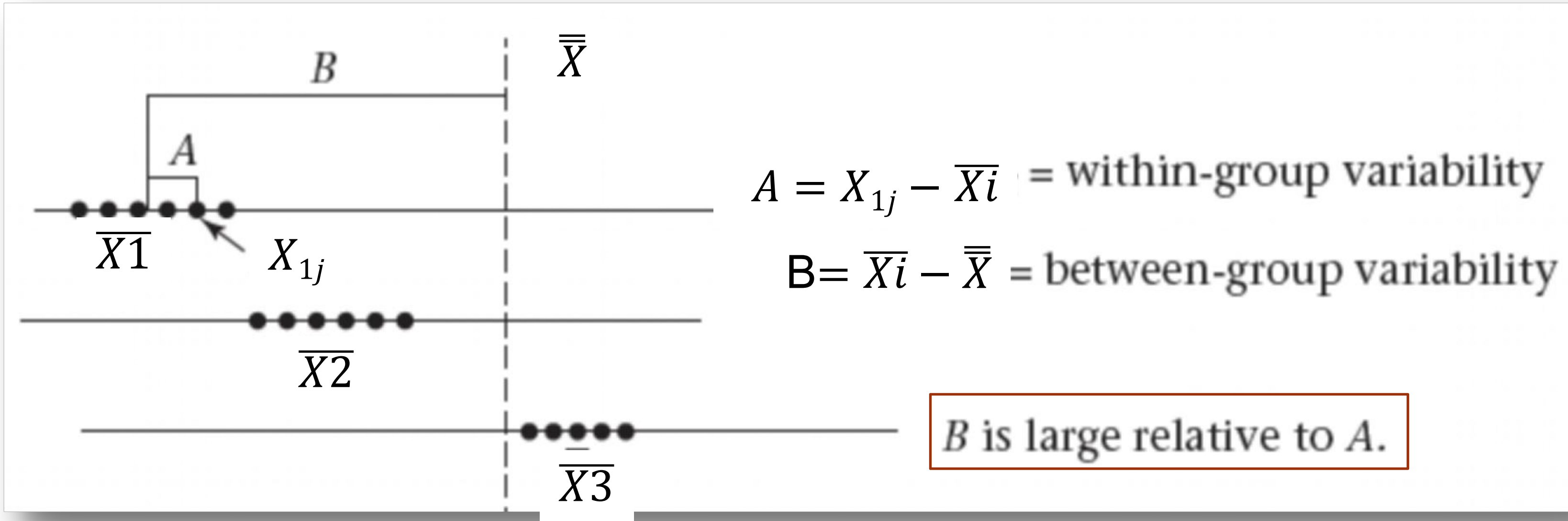


Figure 1.2

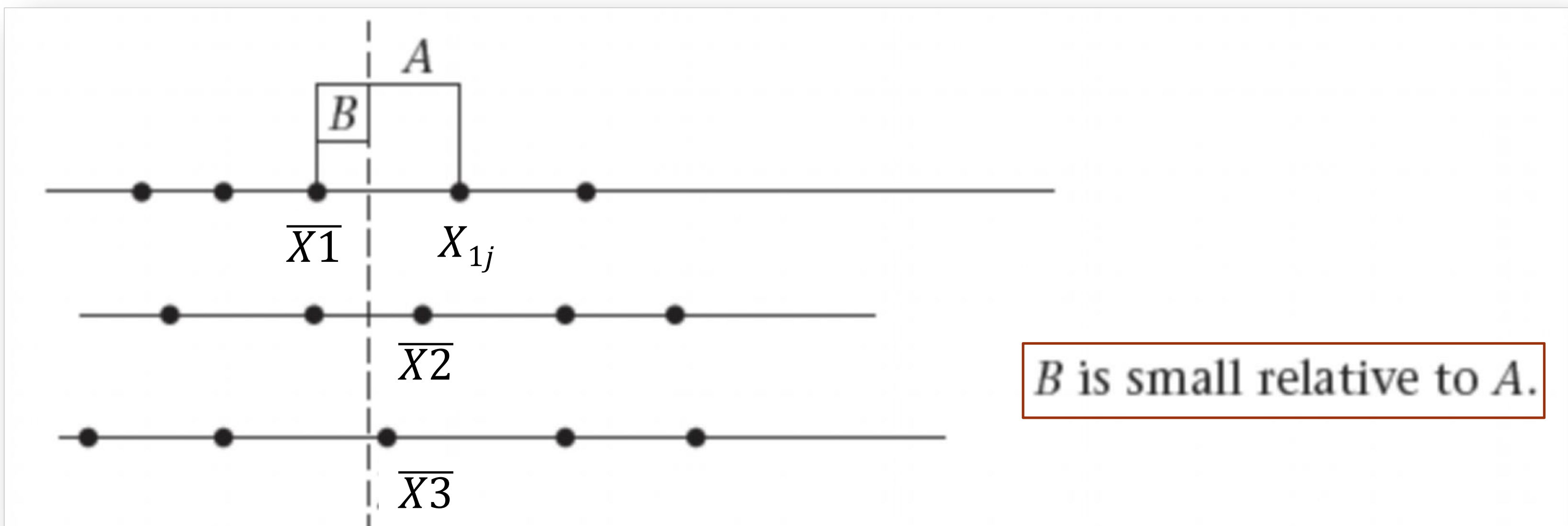


Figure 1.3

Definition 1. Between Mean Square is defined by

$$\mathrm{MSB} = \frac{\mathrm{SSB}}{k-1}.$$

Definition 2. Within Mean Square is defined by

$$\mathrm{MSW} = \frac{\mathrm{SSW}}{n-k}.$$

How to measure relative large/small?

- **Calculate the ratio**

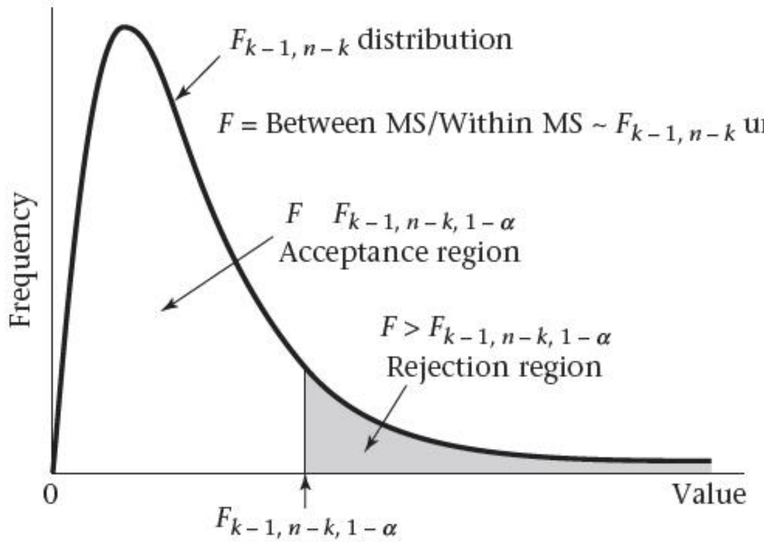
$$\frac{(n-1)*S^2}{\sigma^2} \sim x^2_{n-1}, \text{ then } \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} \sim x^2_{n-1}$$

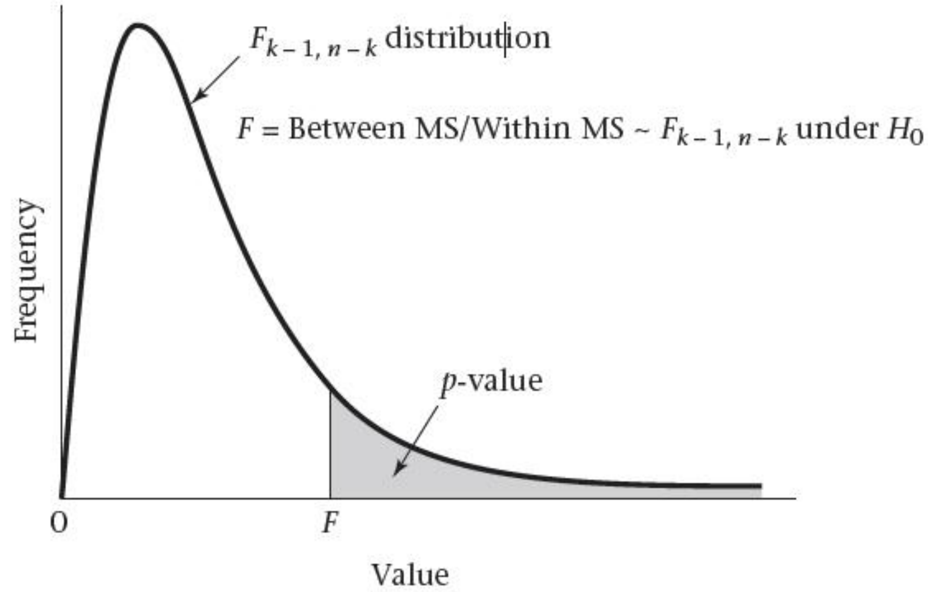
$$\frac{MSB}{MSW} = \frac{SSB/k-1}{SSW/n-k} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{ni}(\bar{X}_i-\bar{\bar{X}})^2 / k-1}{\sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij}-\bar{X}_i)^2 / n-k}$$

If divided by the same σ^2

$$\frac{MSB}{MSW} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{ni}(\bar{X}_i-\bar{\bar{X}})^2 / \sigma^2(k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij}-\bar{X}_i)^2 / \sigma^2(n-k)} = \frac{\chi^2_{k-1}/k-1}{\chi^2_{n-k}/n-k} \sim F_{(k-1,n-k)}$$

$$F = \frac{MSB}{MSW} = \frac{effect + random\ \text{error}}{random\ \text{error}}$$





If $F > F_{k-1,n-k,1-\alpha}$ then reject H_0
If $F \le F_{k-1,n-k,1-\alpha}$ then accept H_0

H0: drugs have no effect

FIGURE 1.4A: REJECTION REGION FOR THE OVERALL F TEST FOR ONE-WAY ANOVA.
FIGURE 1.4B: THE EXACT P-VALUE FOR THE OVERALL F TEST FOR ONE-WAY ANOVA.

(2) Understand in a mathematic/GLM way

Example 1.1:

Drug A	Drug B	Drug C
3	5	5
2	3	6
1	4	7
$\overline{X1}$=2	$\overline{X2}$=4	$\overline{X3}$=6

BRIEF INTRODUCTION OF GLM

- *generalized linear model*

Linear Regression model is not suitable if,

•The relationship between X and y is not linear.

•Variance of errors in y (commonly called as Homoscedasticity in Linear Regression), is not constant, and varies with X.

•Response variable is not continuous, but discrete/categorical.

COMPONENTS OF GLM $Y_I \sim N(X_I^T \beta, \sigma^2)$

- Random Component
- Systematic Component: β0+β1x1+ β2x2 η or $g(\mu)$
- Link function,

$\eta = g(E(Y_i)) = E(Y_i)$ for classical regression

$$\eta = \log(\frac{\pi}{1-\pi}) = \text{logit}(\pi)$$

ASSUMPTIONS OF GLM

- Data should be independent and random
- The response variable y does not need to be normally distributed, but the distribution is from an exponential family (e.g. binomial, Poisson, multinomial, normal)
- The original response variable need not have a linear relationship with the independent variables, but the transformed response variable (through the link function) is linearly dependent on the independent variables

ASSUMPTIONS OF GLM

- Explanatory variables can be nonlinear transformations of some original variables.
- The homogeneity of variance does NOT need to be satisfied. In fact, it is not even possible in many cases given the model structure.
- Errors need to be independent but NOT normally distributed.
- Parameter estimation uses maximum likelihood estimation (MLE) rather than

One-Way ANOVA Fixed-Effects Model

Definition

- **(One-Way ANOVA Fixed-Effects Model):** Suppose there are k groups with n_i observations in the ith group ($i = 1, \dots, k$). The jth observation ($j = 1, \dots, n_i$) in the ith group will be denoted by y_{ij}. Let's assume:

$$\boldsymbol{X_{ij} = \mu + \theta_i + e_{ij},}$$

- where μ is a constant, θ_i is a constant specific to the ith group, and e_{ij} is an error term, which is *normally distributed* with mean 0 and variance σ^2. Thus, a typical observation X_{ij} from the ith group is normally distributed: $\boldsymbol{X_{ij} \sim N(\mu + \theta_i, \sigma^2)}$.

μ, the underlying mean of all groups taken together.

θ_i, the difference between the mean of the ith group and the overall mean.

e_{ij}, random error about the mean $\mu + \theta_i$ for an individual observation from the ith group.

NECESSARY CONSTRAINTS

- It is not possible to estimate both the overall constant μ as well as all the k group specific constants θ_i.

- The reason is that we only have *k observed mean values* for the k groups, which *are insufficient to estimate* $k + 1$ *parameters*.

- As a result, we need to *constrain the parameters* so that *only k parameters will be estimated*. Some typical constraints are: (1) $\theta_1 + \ldots + \theta_k = 0$, or (2) $\theta_k = 0$.

$X_{ij} \sim N(\mu + \theta_i, \sigma^2)$, i=1,2,…,k;
X_{ij} are indpendent, j=1,2,….ni;
μ and σ are not known.

$X_{ij} = \mu_i + \varepsilon_{ij} = \mu + \boxed{\theta_i} + \varepsilon_{ij}$,
$\varepsilon_{ij} \sim N(0, \sigma^2)$ and are indipenent,
μ and σ are not known.

θ_i represents the effects of different drugs(factors).

$H_0: \mu_1 = \ldots = \mu_k$; H_1: not all of the μ_i are identical.

$H_0: \theta_1 = \ldots = \theta_k = 0$; H_1: not all of the θ_i are zero.

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij} - \bar{\bar{X}})^2 = \sum_{i=1}^{k}\sum_{j=1}^{ni}(\boxed{X_{ij} - \bar{X}_i} + \boxed{\bar{X}_i - \bar{\bar{X}}})^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij} - \bar{X}_i)^2 + \sum_{i=1}^{k}\sum_{j=1}^{ni}(\bar{X}_i - \bar{\bar{X}})^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{\bar{X}})$$

The cross-product term can be shown to be zero

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij} - \bar{X}_i)^2 + \sum_{i=1}^{k}\sum_{j=1}^{ni}(\bar{X}_i - \bar{\bar{X}})^2 = SSW + SSB$$

$$SSW=\sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij}-\bar{X}_i)^2 \qquad SSB=\sum_{i=1}^{k}\sum_{j=1}^{ni}(\bar{X}_i-\bar{\bar{X}})^2=\sum_{j=1}^{ni}n_i(\bar{X}_i-\bar{\bar{X}})^2$$

$X_{ij}\sim N(\mu+\theta_i,\sigma^2)$, $\bar{x}_i\sim N(\mu+\theta_i,\frac{\sigma^2}{ni})$, $\bar{\bar{x}}\sim N(\mu,\frac{\sigma^2}{n})$, and $\bar{x}_i$ *are independent*

Then $X_{ij}-\bar{X}_i\sim N(0,\frac{(n_i-1)\sigma^2}{n_i})$, $\mathrm{E}\left[\left(X_{ij}-\bar{X}_i\right)^2\right]=\frac{(n_i-1)\sigma^2}{n_i}$

$$\mathrm{E}\left(SSW\right)=E\left[\sum_{i=1}^{k}\sum_{j=1}^{ni}(X_{ij}-\bar{X}_i)^2\right]=\sum_{i=1}^{k}\sum_{j=1}^{ni}\frac{(n_i-1)\sigma^2}{n_i}=\boxed{(n-k)\sigma^2}$$

Similarly

$$\bar{X}_i-\bar{\bar{X}}\sim N(\theta_i,\frac{(n-n_i)\sigma^2}{nn_i}),\ \mathrm{E}\left[\left(\bar{X}_i-\bar{\bar{X}}\right)^2\right]=\theta_i^2+\frac{(n-n_i)\sigma^2}{nn_i}$$

$$\mathrm{E}\left(SSB\right)=E\left[\sum_{i=1}^{k}n_i(X_{ij}-\bar{X}_i)^2\right]=\boxed{(k-1)\sigma^2+\sum_{i=1}^{k}n_i\theta_i^2}$$

$$\mathrm{E}(SSW)=(n-k)\sigma^2,\quad E\left(\frac{SSW}{n-k}\right)=\sigma^2$$

$$\mathrm{E}(SSB)=(k-1)\sigma^2+\sum_{i=1}^{k}n_i\theta_i^2,\quad \mathrm{E}\left(\frac{SSB}{k-1}\right)=\sigma^2+\frac{1}{k-1}\sum_{i=1}^{k}n_i\theta_i^2\geq\sigma^2$$

if H0 is ture, $\dfrac{SSW}{n-k}$ and $\dfrac{SSB}{k-1}$ are unbiased estimations of σ^2

$$\frac{SSW}{n-k}\approx\frac{SSB}{k-1}$$

If H0 is false $\dfrac{SSW}{n-k}$ will be significantly smaller than $\dfrac{SSB}{k-1}$

$$\frac{MSB}{MSW}\ will\ \text{significantly greater than 1}$$

Display of One-Way ANOVA Results

One-Way ANOVA

Source	DF	Sum of Squares	Mean Squares	F statistic	P value
Between	$k-1$	SSB	$\mathrm{MSB} = \dfrac{\mathrm{SSB}}{k-1}$	$\dfrac{\mathrm{MSB}}{\mathrm{MSW}}$	$\Pr(F_{k-1,n-k} > F)$
Within	$n-k$	SSW	$\mathrm{MSW} = \dfrac{\mathrm{SSW}}{n-k}$		
Total	$n-1$	SSB+SSW			

One-Way ANOVA F-Test Assumptions

1. Randomness & Independence of Errors
2. Normality
   - Populations (for each condition) are Normally Distributed
3. Homogeneity of Variance (*homoscedasticity*)
   - Populations (for each condition) have Equal Variances

2. ONE-WAY ANOVA FOR CRD (COMPLETELY RANDOM DESIGN)

EXAMPLE 2.1

To study the relationship between birth weight and smoking status of mother during pregnancy, one research group performed a prospective study. According to maternal smoking habits, the newborns were divided into four groups.

- Non-smoker (group=1);
- Former smoker (group=2);
- Smoke but less than 1 packet of cigarette (group=3);
- Smoke great than 1 packet of cigarette (group=4).

Test whether the means of newborn weight are equal in the 4 groups.

Table 2.1. Newborn weight according to maternal smoking status

group=1	group=2	group=3	group=4
3.4	2.6	2.7	2.8
2.8	3.3	2.8	3.1
3.1	3.7	2.6	2.6
3.4	3.2	2.1	2.2
4.2	3.5	3.8	2.8
3.8	3.3	3.3	3.2
3.5	3.4	2.8	2.6
3.3	3	2.7	2.5
3.6	3.1	2.9	
3.2	3.2	2.6	
3.7			

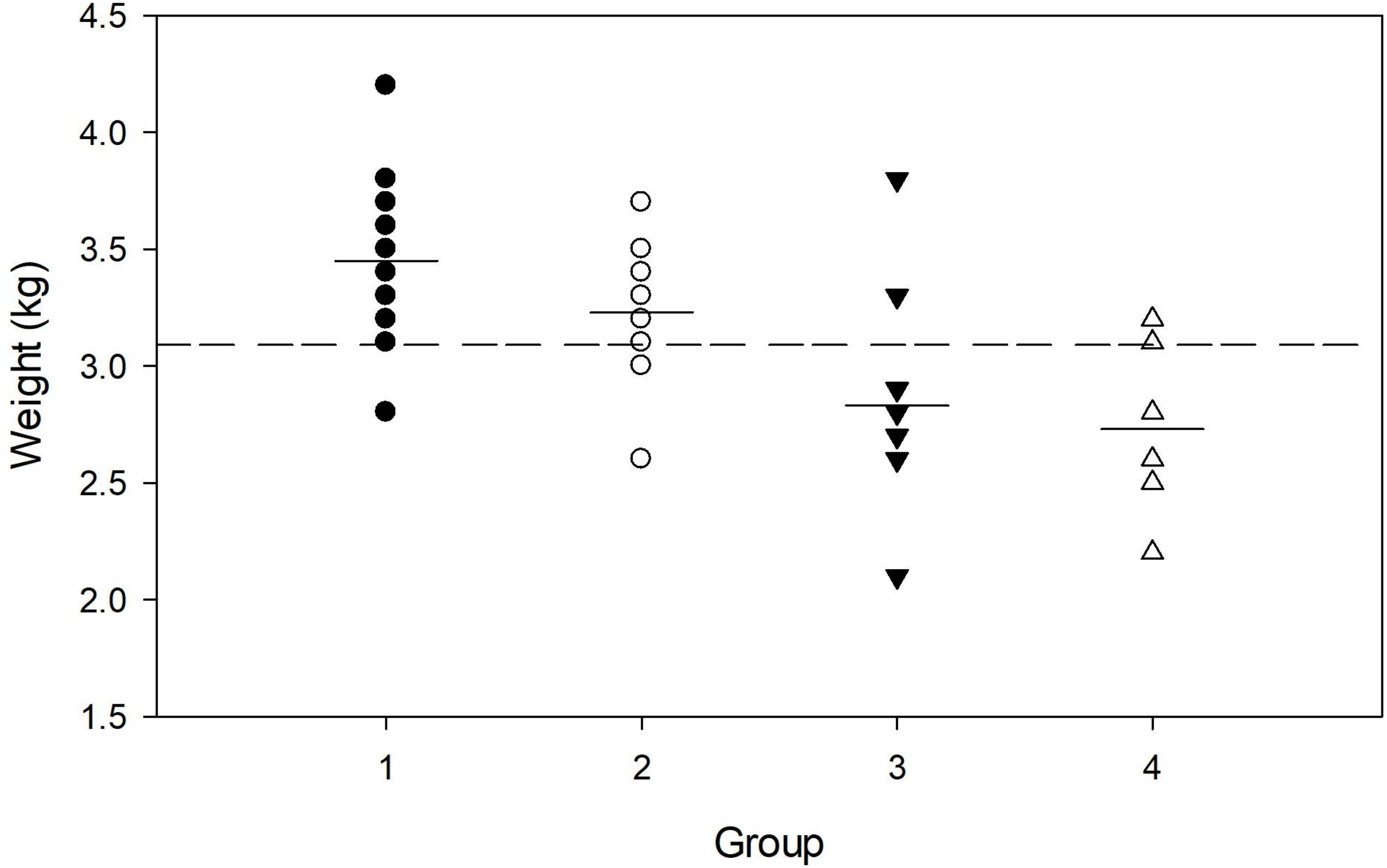


Figure 2.1. The newborn weight according to their maternal smoking status

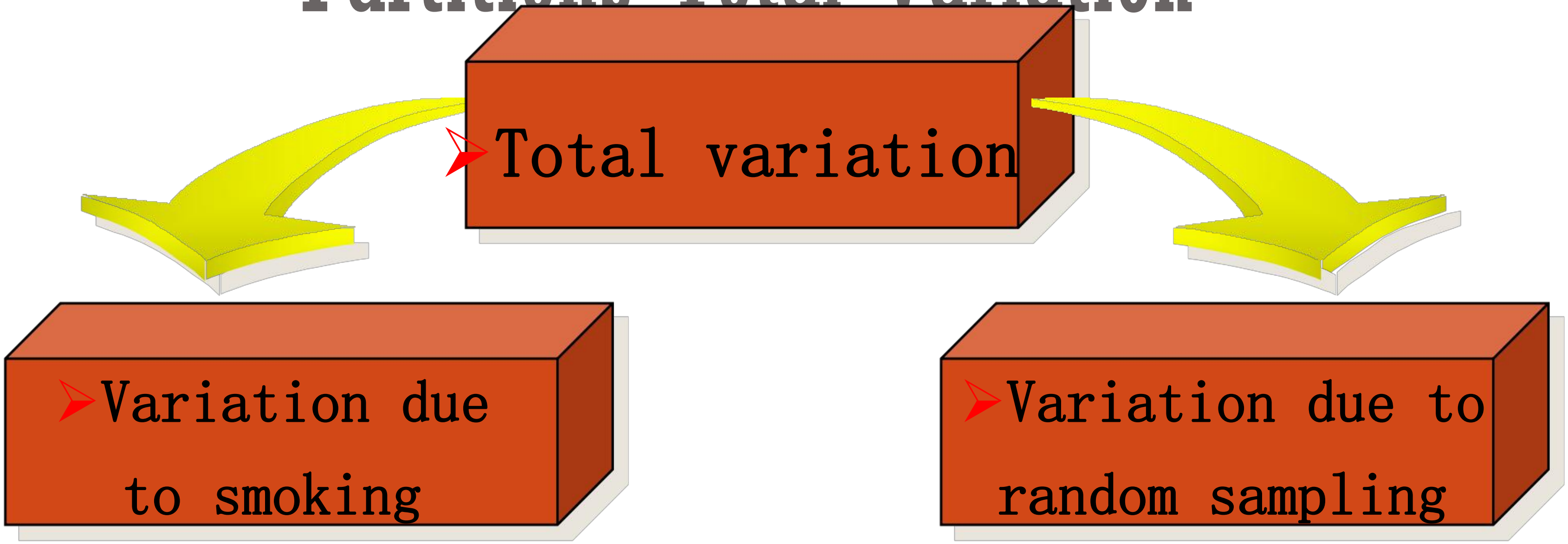
(1) STATE STATISTICAL HYPOTHESES

- Null hypothesis: The (population) means of all groups are equal

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

-
- H_1: There is at least one mean difference among the populations (The pop. Means are not all equal)

One-Way ANOVA Partitions Total Variation



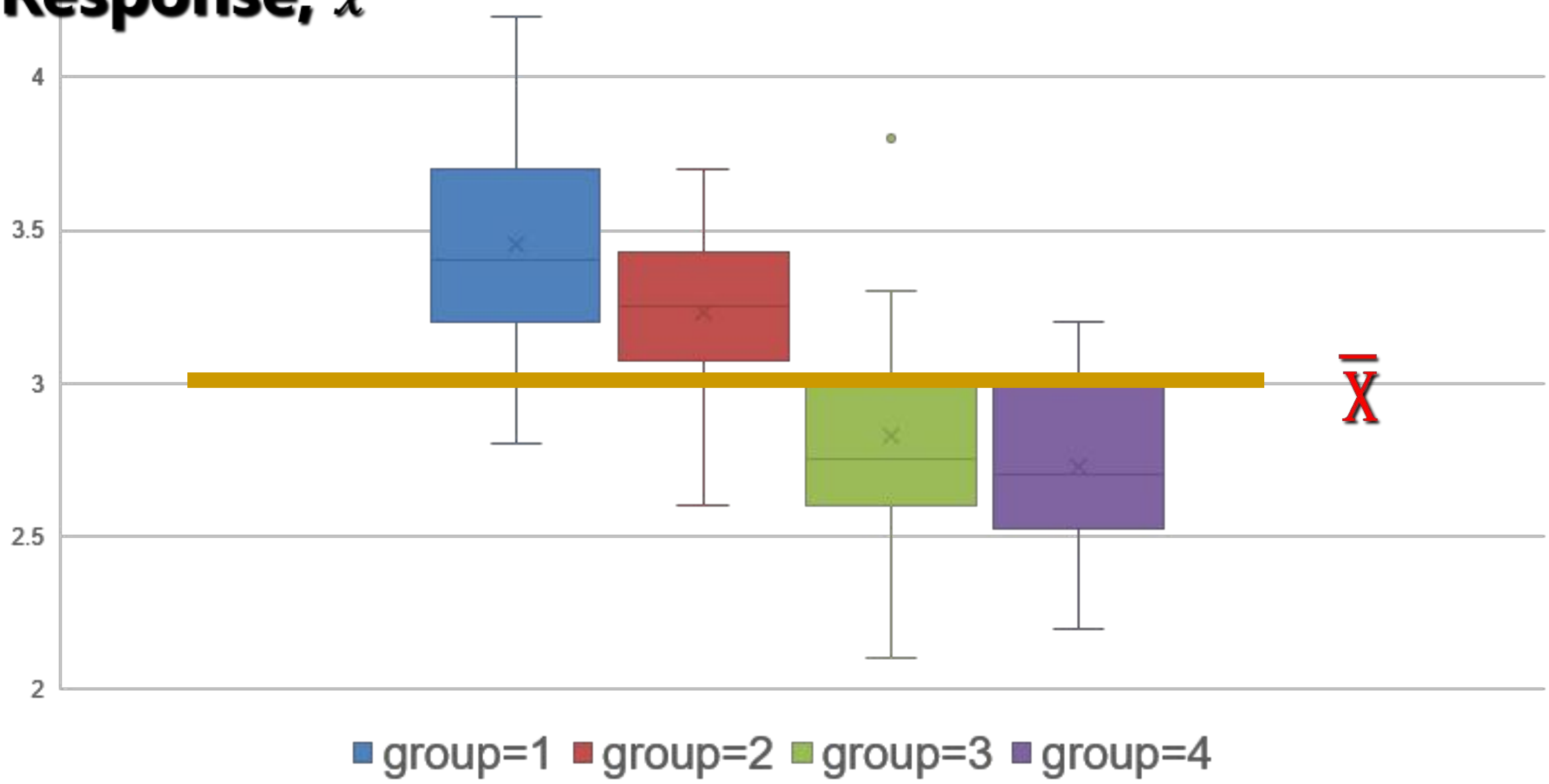
Sum of Squares Among
Sum of Squares Between (SSB)
Sum of Squares Treatment
Among Groups Variation

Sum of Squares Within (SSW)
Sum of Squares Error
Within Groups Variation

TOTAL VARIATION

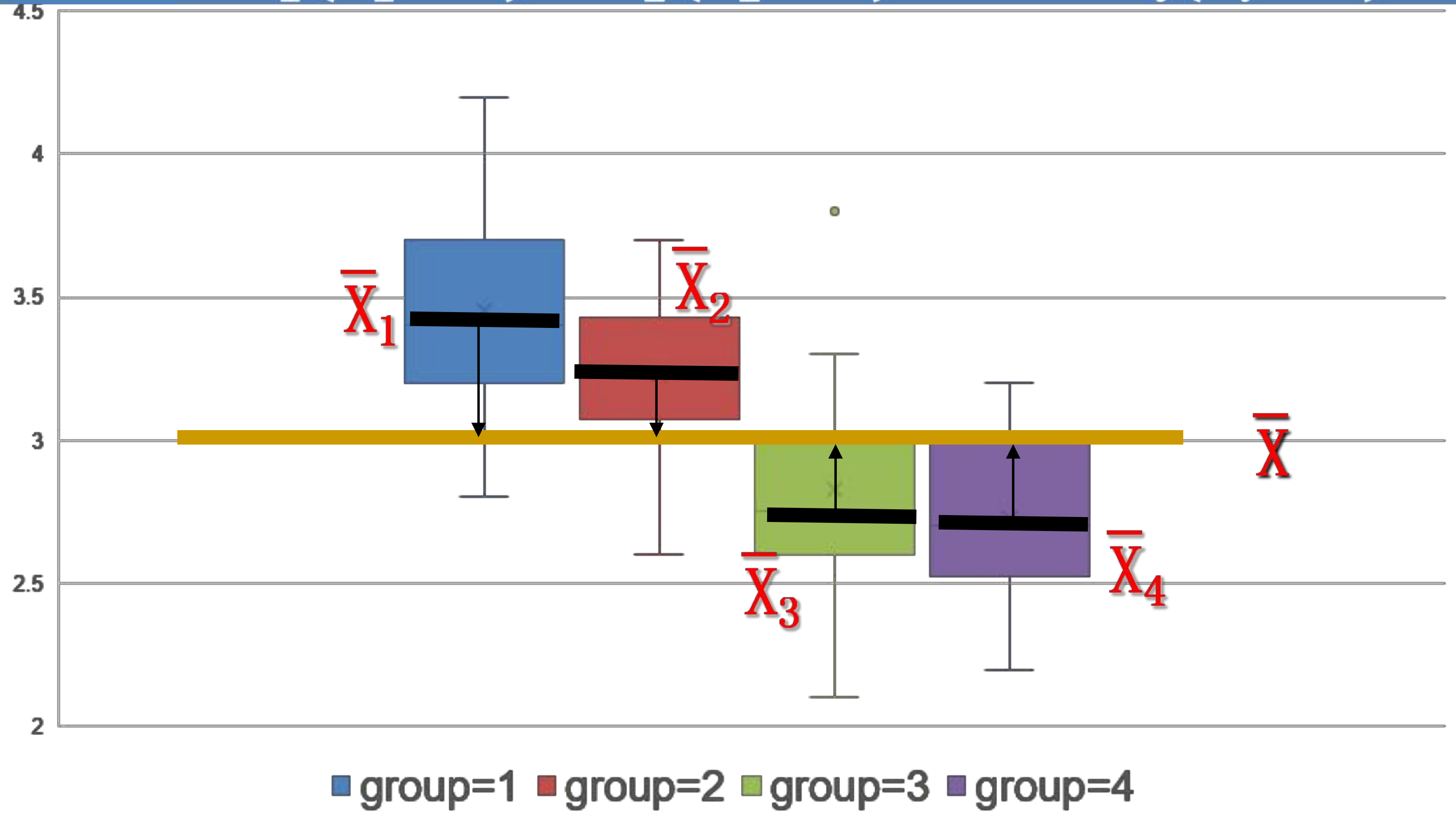
$$SS(total) = (x_{11} - \bar{x})^2 + (x_{21} - \bar{x})^2 + \ldots\ldots + (x_{ij} - \bar{x})^2$$

Response, x



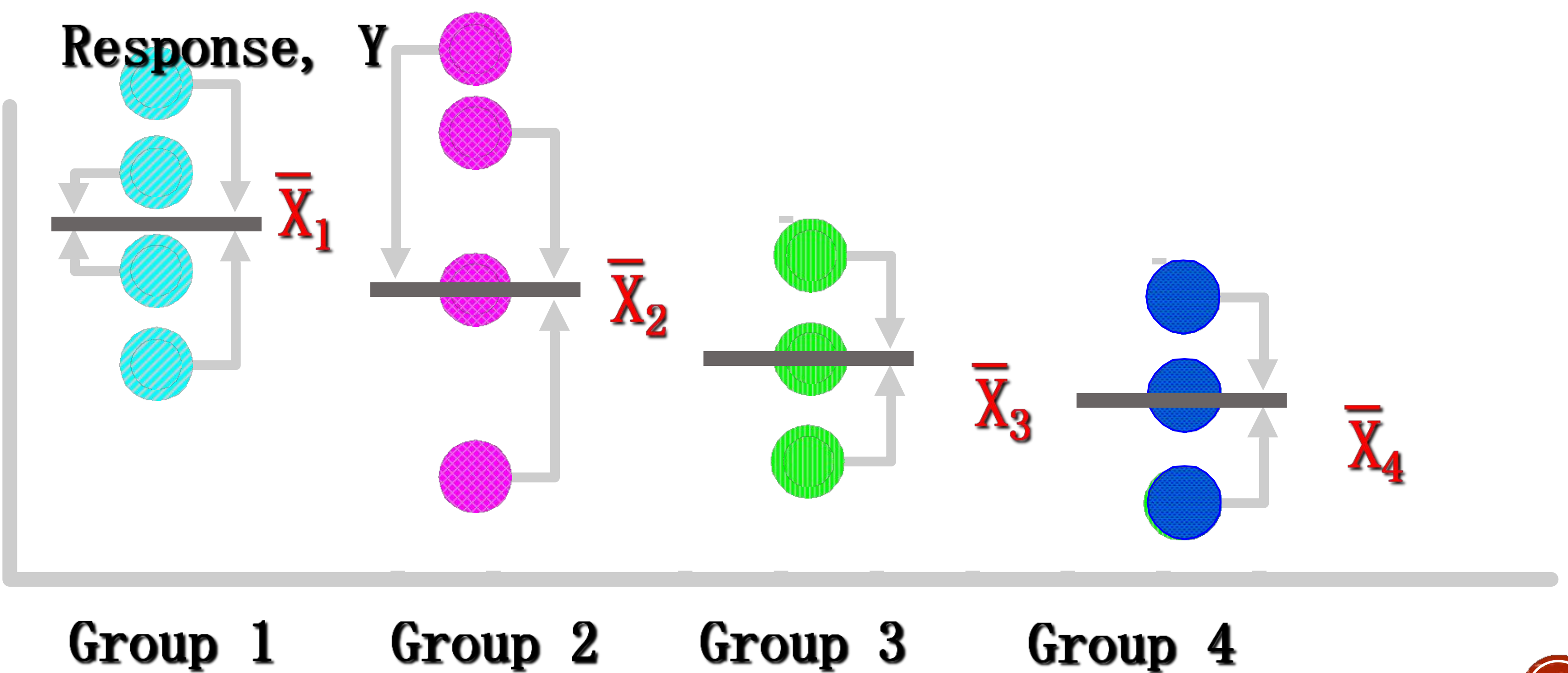
TREATMENT VARIATION (BETWEEN GROUP VARIATION)

SSB $= n_1(\overline{x_1} - \bar{x})^2 + n_2(\overline{x_2} - \bar{x})^2 + \ldots\ldots + n_i(\overline{x_i} - \bar{x})^2$



WITHIN GROUP VARIATION

$$SSW = (x_{11} - \overline{x_1})^2 + (x_{21} - \overline{x_1})^2 + \ldots\ldots + (x_{ij} - \overline{x_i})^2$$



SHORT COMPUTATIONAL FORM

① Total Variation

$$SST = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}\right)^2 = \sum_{i=1}^{g}\sum_{j=1}^{n_i} x_{ij}^2 - C,\quad \mathrm{df_t} = N-1$$

$$C = \left(\sum_{i=1}^{g}\sum_{j=1}^{n_i} x_{ij}\right)^2 / N$$

② Variation between groups

$$SSB = \sum_{i=1}^{g} n_i(\bar{x}_i - \bar{x})^2 = \sum_{i=1}^{g}\frac{(\sum_{j=1}^{n_i} x_{ij})^2}{n_i} - C,\quad df_B = g-1$$

③ Variation within groups

$$SSW = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)^2,\quad \mathrm{df_w} = \sum_{i}(n_i - 1) = N - g$$

N=total sample size

g=# of treatment (level)

MEAN SQUARES AND F-RATIO

$$MS_{between} = s^2_{between} = \frac{SS_{between}}{df_{between}}$$

$$MS_{within} = s^2_{within} = \frac{SS_{within}}{df_{within}}$$

$$F = \frac{s^2_{between}}{s^2_{within}} = \frac{MS_{between}}{MS_{within}}$$

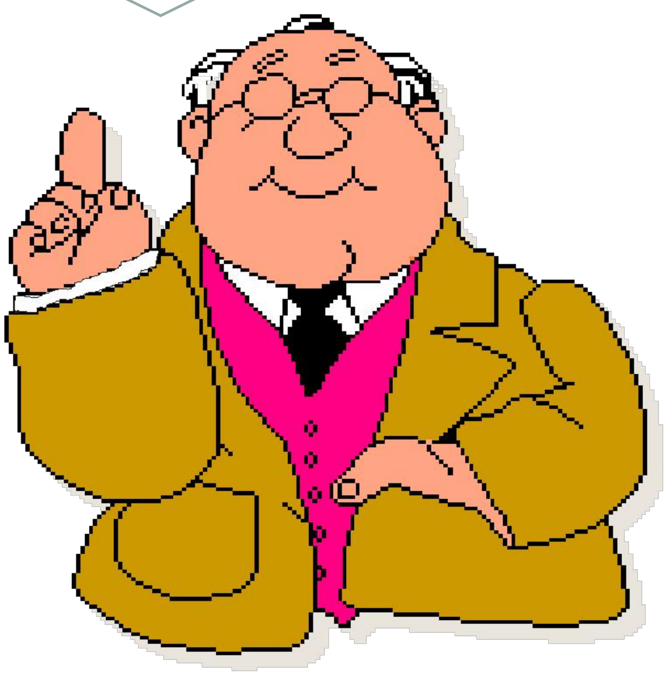
(2) COMPUTE A TEST STATISTIC

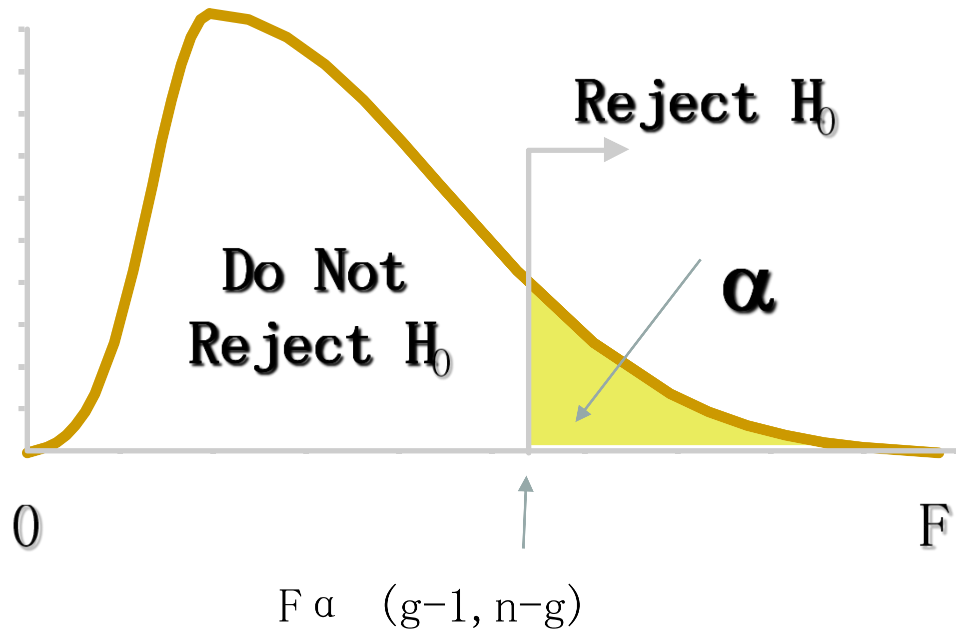
Table 2.2 ANOVA table

Source	df	SS	MS	F	P
Total	38	8.18			
Between group	3	3.40	1.13	8.29	
Within group	35	4.78	0.14		

ONE-WAY ANOVA F-TEST

If means are equal, $F = MSB\ /\ MSW \approx 1$. Only reject large F!





Always One-Tail!

(3) OBTAIN THE CRITICAL VALUE

df_{between} (numerator)

df_{within} (denominator)

df of Denominator	df freedom: Numerator		
	1	2	3
34			2.88
35			**2.87**

(4)DECISION RULE

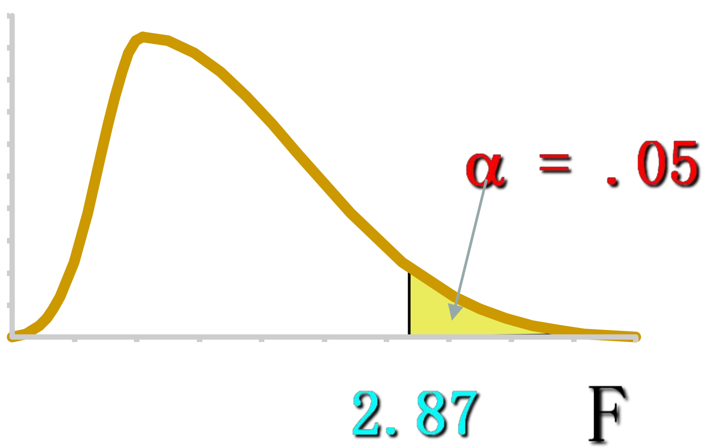
H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

H_a: **Not All Equal**

α = **.05**

ν_1 = **3** ν_2 = **35**

Critical Value(s):



Test Statistic:

$$F = \frac{MSB}{MSW} = \frac{3.40/3}{4.78/35} = 8.29$$

Decision:

Reject at α = .05

Conclusion:

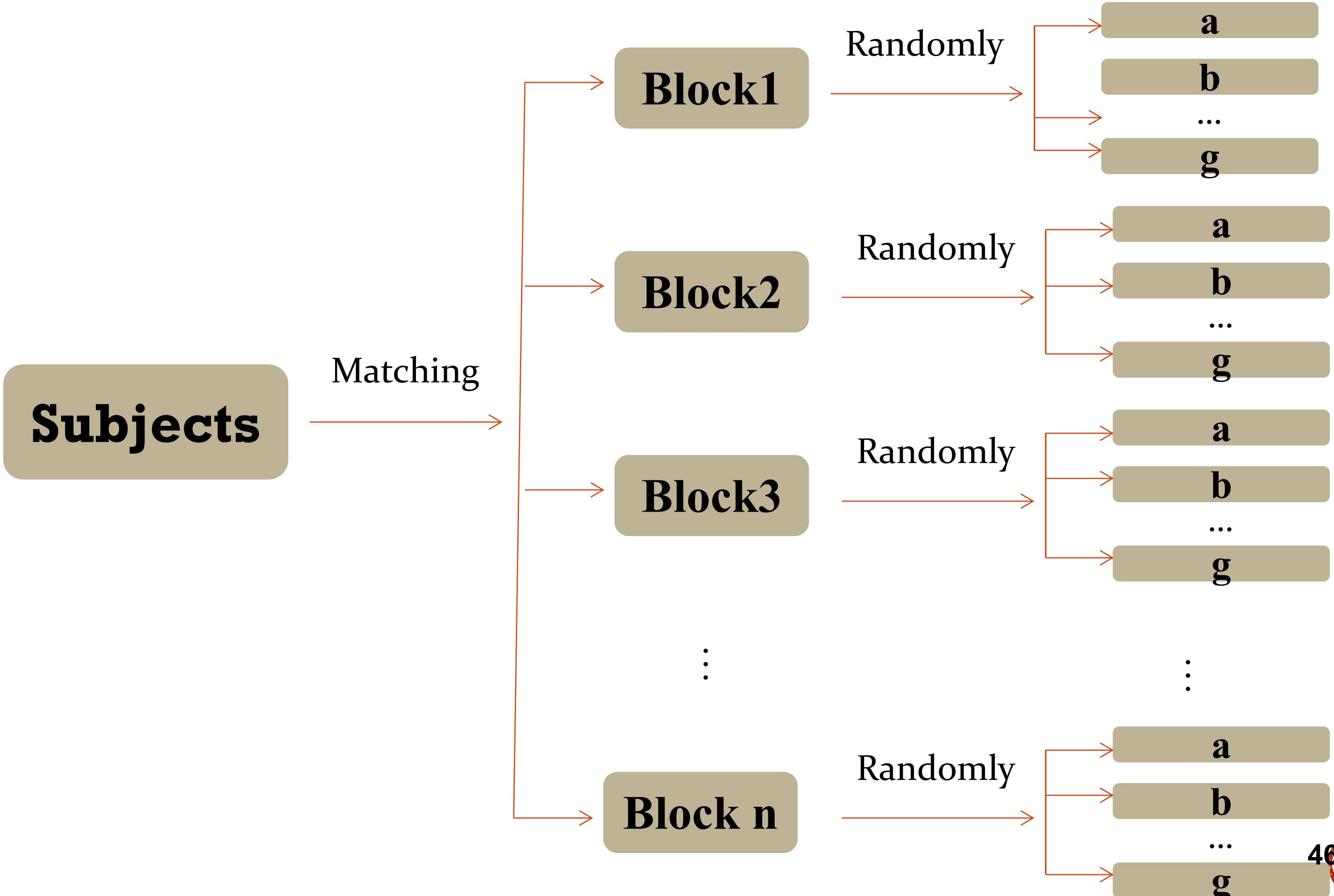
There Is Evidence Pop. Means Are Different

SUMMARY

The final goal for the ANOVA is an *F*-ratio	$F = \dfrac{\text{Variance between treatments}}{\text{Variance within treatments}}$
Each variance in the *F*-ratio is computed as *SS*/*df*	$\text{Variance between treatments} = \dfrac{SS \text{ between}}{df \text{ between}}$ $\text{Variance within treatments} = \dfrac{SS \text{ within}}{df \text{ within}}$
To obtain each of the *SS* and *df* values, the total variability is analyzed into the two components	*SS* total: *SS* between, *SS* within *df* total: *df* between, *df* within

3. ANOVA FOR RBD (RANDOMIZED BLOCK DESIGN)

Randomized Block Design



Randomized Block F-Test Assumptions

1. Normality
   - Probability Distribution of each Block-Treatment combination is Normal

2. Homogeneity of Variance
   - Probability Distributions of all Block-Treatment combinations have Equal Variances

Example 3.1: Nine individuals were allocated to 3 different blocks according to their gender and BMI, then 3 individuals within a certain block were randomly assigned to take different drugs, and then the serum levels of a special protein were tested.

	Drug A	Drug B	Drug C	
Block 1	3	5	5	MB1=4.333
Block2	2	3	6	MB2=3.667
Block3	1	4	7	MB3=4
	$\overline{X1}$=2	$\overline{X2}$=4	$\overline{X3}$=6	

The ground (overall) mean: $\overline{\overline{X}}$=4

random error

Not every value within the same drug group are identical

Not every value within the same block are identical

Why?

effect of different drugs

group means are not identical

	Drug A	Drug B	Drug C	
Block 1	3	5	5	MB1=4.333
Block2	2	3	6	MB2=3.667
Block3	1	4	7	MB3=4
	$\overline{X1}=2$	$\overline{X2}=4$	$\overline{X3}=6$	

$SST=(3-4)^2+(2-4)^2+(1-4)^2+(5-4)^2+(3-4)^2+(4-4)^2+(5-4)^2+(6-4)^2+(7-4)^2=30$

Df=3×3-1

$SS_{treatment}=(2-4)^2+(2-4)^2+(2-4)^2+(4-4)^2+(4-4)^2+(4-4)^2+(6-4)^2+(6-4)^2+(6-4)^2=24$

Df=3-1

$SS_{block}=3\times(4.33-4)^2+3\times(3.667-4)^2+3\times(4-4)^2=0.659$

Df=3-1=block-1

$SS_{error}=SST-SS_{treatment}-SS_{block}=5.341$

Df=（block-1）*（treatment-1）=4

$SST=SS_{treatment}+SS_{block}+SS_{error}$

VARIATION DECOMPOSITION

The previous within group variance

$$SS_{total} = SS_{treatment} + \boxed{SS_{block} + SS_{error}}$$

$$\nu_{total} = \nu_{treatment} + \nu_{block} + \nu_{error}$$

EXAMPLE 3.1

- To explore the protective effect of hexadecadrol on animal models of acute lung injury, 36 SD mice were divided into 12 blocks according to gender and weight, and then 3 mice in the same block were randomly assigned to control, injury and hexadecadrol groups. 24h later, the total albumin levels were tested and showed in the following table. Are the albumin levels identical in the three groups?

Table 3.1. Total albumin levels (g/L) in different treatment groups

Block	Control	Injury	hexadecadrol
1	0.36	1.48	0.30
2	0.28	1.42	0.32
3	0.26	1.33	0.29
4	0.25	1.48	0.16
5	0.36	1.26	0.35
6	0.31	1.53	0.43
7	0.33	1.40	0.31
8	0.28	1.30	0.13
9	0.35	1.58	0.33
10	0.41	1.24	0.32
11	0.49	1.47	0.26
12	0.27	1.32	0.26

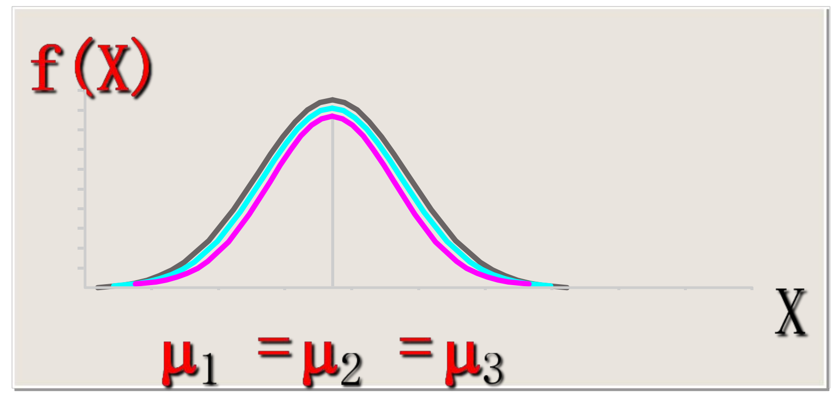
Randomized Block F-Test step1: Hypotheses

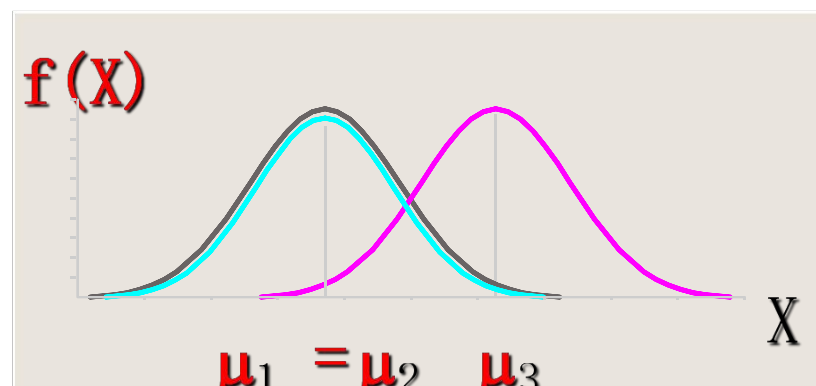
H_0: $\mu_1 = \mu_2 = \ldots = \mu_p$

- **All Population Means are Equal**
- **No Treatment Effect**

H_a: Not All μ_j Are Equal

- **At Least 1 Pop. Mean is Different**
- **Treatment Effect**
- **$\mu_1 \neq \mu_2 \neq \ldots \neq \mu_p$ Is wrong**





RANDOMIZED BLOCK F-TEST
STEP 2: CALCULATION TEST STATISTIC

$$SS_{\text{total}} = \sum\sum x_{ij}^2 - C = 9.8109$$

$$\nu_{\text{total}} = N - 1 = 35$$

$$SS_{\text{treatment}} = \sum_{i=1}^{g} \frac{\left(\sum_{j=1}^{b} x_{ij}\right)^2}{b} - C = 9.5511$$

$$\nu_{\text{treatment}} = g - 1 = 2$$

$$SS_{\text{block}} = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{g} x_{ij}\right)^2}{g} - C = 0.1138$$

$$\nu_{\text{block}} = b - 1 = 11$$

$$SS_{\text{error}} = SS_{\text{total}} \text{-} SS_{\text{treatment}} \text{-} SS_{\text{block}} = 0.1460$$

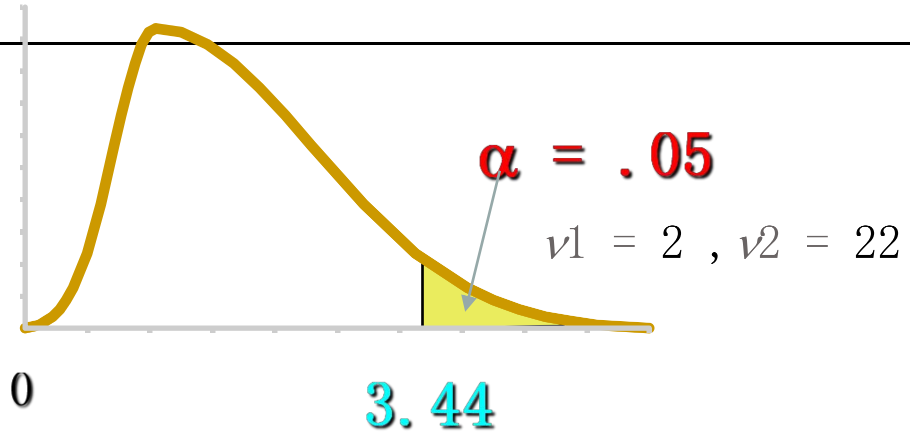
$$\nu_{\text{error}} = \nu_{\text{total}} \text{-} \nu_{\text{treatment}} \text{-} \nu_{\text{block}} = 22$$

Randomized Block F-Test Test Statistic

- 1. Test Statistic
  - $F = MST / MSE$
    - MST Is Mean Square for Treatment
    - MSE Is Mean Square for Error
- 2. Degrees of Freedom
  - $\nu_1 = g - 1$
  - $\nu_2 = (b-1)(g-1)$
    - g = # Treatments, b = # Blocks, n = Total Sample Size

Randomized Block F-Test
Step 3-P Value and conclution

Source	df	SS	MS	F	P
Total	35	9.8109			
Treatment	2	9.5511	4.7756	719.80	<0.05
Block	11	0.1138	0.0103	1.56	>0.05
Error	22	0.1460	0.0066		



STEP 4. DECISION RULE

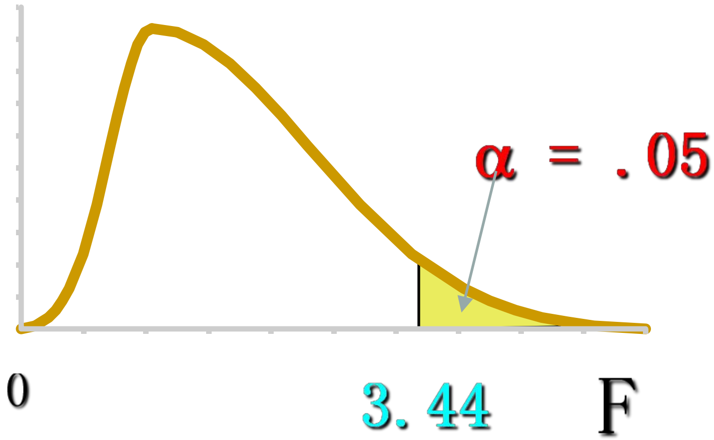
H_0: $\mu_1 = \mu_2 = \mu_3$

H_a: **Not All Equal**

α = **.05**

$\nu_1 = 2 \quad \nu_2 = 22$

Critical Value(s):



Test Statistic:

$$F = \frac{MST}{MSE} = \frac{4.7756}{0.0066} = 719980$$

Decision:

Reject at $\alpha = .05$

Conclusion:

There Is treatment effect