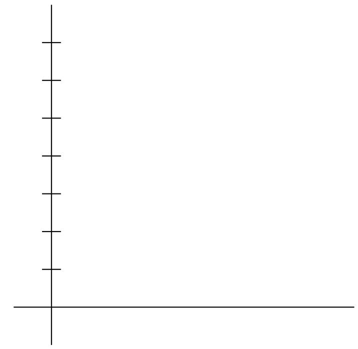
9.1 Review: One-Way ANOVA
9.2 Pairwise Comparisons in One-Factor Experiments
9.3 Two-Factor Experiments (part 1 of 2)



9.1 Review: One-Way ANOVA

e.g. Make a quick one-way ANOVA table for these data to test $H_0 : \mu_1 = \mu_2 = \mu_3$:

Sample	Data	$\bar{x}_{i.}$
1	0 1 2	
2	1 2 3	
3	5 6 7	
	Overall $\bar{x}_{..} =$	

Source	DF	SS	MS	F	P
Treatment	$I-1$	SSTr $= \sum_{i=1}^{I} J_i(\bar{X}_{i.} - \bar{X}_{..})^2$	MSTr $= \frac{SSTr}{I-1}$	$F = \frac{\text{MSTr}}{\text{MSE}}$	$P(F_{I-1,N-I} > \_\_\_\_)$
Error	$N-I$	SSE $= \sum_{i=1}^{I} \sum_{j=1}^{J_i}(X_{ij} - \bar{X}_{i.})^2$	MSE $= \frac{SSE}{N-I}$ ($= s^2$)		
Total	$N-1$	SST $= \sum_{i=1}^{I} \sum_{j=1}^{J_i}(X_{ij} - \bar{X}_{..})^2$ $=$ SSTr $+$ SSE			

Conclusion: ____________________________

9.2 Pairwise Comparisons in One-Factor Experiments

A problem:

- One-way ANOVA tests $H_0 : \mu_1 = \cdots = \mu_I$ (for I samples), but rejecting H_0 doesn't tell __________________________ are different.
- If we calculate a level-$(1-\alpha)$ confidence interval for the difference of each pair of means, $\mu_i - \mu_j$, the probability of ______ intervals containing their respective differences ___________________ is usually ___________________

The Tukey-Kramer Method of Multiple Comparisons

- Experts say the Tukey-Kramer level-$(1-\alpha)$ ____________ *confidence intervals* for $\mu_i - \mu_j$ are

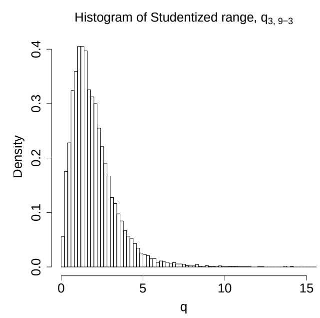
$$\bar{X}_{i.} - \bar{X}_{j.} \pm q_{I,N-I,\alpha} \sqrt{\frac{\text{MSE}}{2}\left(\frac{1}{J_i} + \frac{1}{J_j}\right)}$$

  where $q_{I,N-I,\alpha}$ cuts off a right tail area α from the __________________________ *distribution* with I and $N-I$ degrees of freedom. (Table A.7, p. 534-535, gives this q in column $\nu_1 =$ ______, row $\nu_2 =$ ______, and subrow ______.)
  We have $(1-\alpha)$ confidence that these intervals contain $\mu_i - \mu_j$ __________________________

- To test all $H_0 : \mu_i - \mu_j = 0$ ____________________, use the test statistics

$$q_{i,j} = \frac{|\bar{X}_{i.} - \bar{X}_{j.}| - 0}{\sqrt{\frac{\text{MSE}}{2}\left(\frac{1}{J_i} + \frac{1}{J_j}\right)}}$$

  The P-value for test i, j is _____________
  To _____________ calculating $\binom{I}{2}$ P-values, here's a _____________:



$$H_0 : \mu_i - \mu_j = 0 \text{ is rejected at level } \alpha \iff q_{i,j} > \text{\_\_\_\_\_\_\_\_\_\_\_\_\_} \text{ (draw)}$$

$$\iff |\bar{X}_{i.} - \bar{X}_{j.}| > q_{I,N-I,\alpha} \sqrt{\frac{\text{MSE}}{2}\left(\frac{1}{J_i} + \frac{1}{J_j}\right)}$$

e.g. Which pairs of population means from the "9.1 Review" example, above, differ at the 5% level?

9.3 Two-Factor Experiments (part 1 of 2)

Example

e.g. (p. 437 #6) A tool's lifetime was studied under three settings for feed rate and three for speed. Four tools were tested under each combination of settings. The lifetimes (in hours) were:

	Speed		
Feed Rate	Slow	Medium	Fast
Light	60.6 57.0 61.4 59.7	57.8 59.4 62.8 58.2	56.5 52.3 58.1 53.9
Medium	51.2 53.1 48.3 51.6	49.6 48.1 49.8 51.1	45.7 48.6 45.0 49.2
Heavy	44.8 46.7 41.9 51.3	46.6 41.4 38.3 37.9	37.2 32.8 39.9 35.9

Does feed rate influence lifetime? Does speed influence lifetime?

Vocabulary

Recall (§9.1) that a *factor* is an ______________ variable in an experiment; its values are ___________.

A *two-factor experiment* has a _____________ *factor* with I levels and a _____________ *factor* with J levels, for a total of _____________ combinations called _____________.

An *experimental* _______ is an object on which measurements are made. _____________ are units assigned to a given treatment. A _____________ design uses the same number _______ of replicates for each treatment. (We analyze only balanced designs.)

In a _____________ or *full factorial design*, observations are taken on _______ of the IJ possible treatments. (Incomplete designs, which lack data for some treatments, are hard to interpret.)

Each treatment represents a _____________ whose mean outcome is μ_{ij}. Replicates corresponding to level i of the row factor and level j of the column factor are denoted $X_{ij1}, \cdots, X_{ijK}$ and represent a ___________________________ from this population.

A *two-way analysis of variance* tests whether $\{\mu_{ij}\}$ are affected by varying the ____________________ or ___________________________ or both.

Parameterization for Two-Way Analysis of Variance

- $\mu_{ij} =$ _____________ mean corresponding to treatment in row _______ and column _______
- $\bar{\mu}_{i.} = \frac{1}{J}\sum_{j=1}^{J}\mu_{ij} =$ average of means in _____________
- $\bar{\mu}_{.j} = \frac{1}{I}\sum_{i=1}^{I}\mu_{ij} =$ average of means in _____________
- $\mu = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\mu_{ij}\left(= \frac{1}{I}\sum_{i=1}^{I}\bar{\mu}_{i.} = \frac{1}{J}\sum_{j=1}^{J}\bar{\mu}_{.j}\right) =$ *population grand mean* = average of _____________ treatment means

  Here's a table showing these (unknown) population means:

	Column level			
Row level	1	$\cdots$	J	Row mean
1	μ_{11}	$\cdots$	μ_{1J}	$\bar{\mu}_{1.}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
I	μ_{I1}	$\cdots$	μ_{IJ}	$\bar{\mu}_{I.}$
Column mean	$\bar{\mu}_{.1}$	$\cdots$	$\bar{\mu}_{.J}$	μ

- Write μ_{ij} as

$$\begin{aligned} \mu_{ij} &= \mu + (\bar{\mu}_{i.} - \mu) + (\bar{\mu}_{.j} - \mu) + (\mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \mu) \\ &= \mu + \alpha_i + \beta_j + \gamma_{ij} \end{aligned}$$

  where

  - $\alpha_i = \bar{\mu}_{i.} - \mu = i^{\text{th}}$ *row effect*, which indicates how much ______________________________ ________ tends to produce outcomes larger or smaller than μ (note: $\sum_{i=1}^{I} \alpha_i =$ _______)
  - $\beta_j = \bar{\mu}_{.j} - \mu = j^{\text{th}}$ *column effect*, which indicates how much ______________________________ ________ tends to produce outcomes larger or smaller than μ (note: $\sum_{j=1}^{J} \beta_j =$ _______)
  - $\gamma_{ij} = \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \mu$ [$= \mu_{ij} - (\mu + \alpha_i + \beta_j)$], the ij ______________, which measures the degree to which the effect of a row (or column) factor depends on the level of the column (or row) factor it is paired with (note: $\sum_{i=1}^{I} \gamma_{ij} =$ _______ and $\sum_{j=1}^{J} \gamma_{ij} =$ _______)

  Main effects consist of the I row effects and the J column effects; they don't include the _____ ________________. If the interactions γ_{ij} are all 0, then $\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$ becomes

$$\mu_{ij} = \mu + \alpha_i + \beta_j$$

  In this case we have an ______________ *model*, such that the treatment mean μ_{ij} is the population grand mean μ plus the amount α_i resulting from using row level i plus the amount β_j resulting from using column level j. (If some interactions are not zero, the combined effect of a row and column can't be determined from the _____________________ alone.)

Using Two-Way ANOVA to Test Hypotheses

Two-Way ANOVA tests

- Does the ______________ model hold? That is, is it plausible that $\gamma_{ij} =$ ____ for all i and j?
  $H_0 : \gamma_{11} = \cdots = \gamma_{IJ} = 0$
- If the additive model holds (H_0 is ______________),
  - is the mean outcome ______________ for all levels of the row factor?
    $H_0 : \alpha_1 = \cdots = \alpha_I = 0$
  - is the mean outcome ______________ for all levels of the column factor?
    $H_0 : \beta_1 = \cdots = \beta_J = 0$

Retain all $\Longrightarrow$ "$\mu_{ij} =$ ____ for all i, j" is plausible $\Longrightarrow$ treatment differences ______________