Exam 2 Stat 502 WD Fall 2016

1. At the end of the school year in each of four randomly selected schools (S), two first-grade classes (C) were randomly selected. Within each class, ten boys and ten girls (G for gender) were randomly selected, half of which were assigned to one of two treatments (T): parental guided reading (E) and control (X).

a. (5 pts) How many students are involved in the study?

n x G x C x S x T = 5 x 2 x 2 x 4 x 2 = 160

160

b. (5pts) In the scope of inference, which factors are random effects (if any) and which factors (if any) are fixed effects?

~~Random: School, Class, Student~~

~~Fixed: Gender~~

Random: Schools (S) and Classes (C)

Fixed: Treatment (T) and Gender (G)

c. (5pts) In the following table, list all the factor sources (i.e. exclude error and total) for this study and identify them as either random or fixed. Note there *might* be more rows than possible sources!

|  |  |
| --- | --- |
| ~~Source~~ | ~~Random or Fixed?~~ |
| ~~School~~ | ~~Random~~ |
| ~~Class (School)~~ | ~~Random~~ |
| ~~Gender (School Class)~~ | ~~Fixed~~ |
| ~~Student (School Class Gender)~~ | ~~Random~~ |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

|  |  |
| --- | --- |
| Source | Random or Fixed? |
| Treatment (T) | Fixed |
| Gender (G) | Fixed |
| School (S) | Random |
| Classes (School), C(S) | Random |
| T x G | Fixed |
| T x S | Random |
| T x C(S) | Random |
| G x S | Random |
| G x C(S) | Random |
| T x G x S | Random |
| T x G x C(S) | Random |

2) A plastic extrusion company wants to compare three particular mixes for its plastic, and is also interested comparing the three plasticizers that they currently use to harden the plastic. Each day they make one run, by first mixing up three large tubs of plastic, one for each of the mixes (M1, M2, M3). The three plasticizers are then added to portions of the tub with bulk mix prior to extrusion into a product. Product strength is measured as the response variable. They make three runs of this experiment and want to use an ANOVA.

We have here:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Plasticizer | Mix | | |
| Run |  | M1 | M2 | M3 |
| 1 | P1 | Y | Y | Y |
|  | P2 | Y | Y | Y |
|  | P3 | Y | Y | Y |
|  |  |  |  |  |
| 2 | P1 | Y | Y | Y |
|  | P2 | Y | Y | Y |
|  | P3 | Y | Y | Y |
|  |  |  |  |  |
| 3 | P1 | Y | Y | Y |
|  | P2 | Y | Y | Y |
|  | P3 | Y | Y | Y |

1. (5 pts) Identify (i.e. name), the treatment design and randomization design involved in this study.

~~Treatment design is 3x3 factorial design.~~

Treatment design: 2-factor factorial with Mix x Plasticizer

Split-Plot in RCBD.

Each run (block) has all combinations of mix x plasticizer levels represented (RCBD). However, the mixes are first established in each run, and then the mix is split into 3 subgroups that subsequently receive a plasticizer method, forming a split plot design within the RCBD structure.

b) (5 pts) Provide the ANOVA table, including columns for source and degrees of freedom.

| **Type 3 Analysis of Variance** | | | |
| --- | --- | --- | --- |
| **Source** | **DF** | **Expected Mean Square** |
| **mix** | 2 | Var(Residual) + 3 Var(block \* mix) + Q(mix, mix \* plasticizer) |
| **plasticizer** | 2 | Var(Residual) + Q(plasticizer, mix \* plasticizer) |
| **mix\*plasticizer** | 4 | Var(Residual) + Q(mix \* plasticizer) |
| **block** | 2 | Var(Residual) + 3 Var(block\* mix) + 12 Var(block) |
| **block\*mix** | 4 | Var(Residual) + 3 Var(block \* mix) |
| **Residual** | 12 | Var(Residual) |

c) (5 pts) How would the F-test for the Mix treatment be computed?

Mix = MSmix / MSblock\*mix

Note that the correct error term for the F test of the treatment applied to whole plots is the block × treatment interaction since blocks (runs) are random effects.

3) In a pilot test of a medical machine, 4 machines were chosen at random from the pilot assembly production to evaluate their performance in assaying triglyceride levels (mg/dl) in serum samples. Eight replicate serum samples were prepared by each technician (chosen at random from a pool of available technicians), and 2 serum samples were randomly assigned to each of the four machines for each technician.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ***Machine*** | | | |
| ***Technician*** | **1** | **2** | **3** | **4** |
| **1** | 142.3, 144.0 | 148.6, 146.9 | 142.9, 147.4 | 133.8. 133.2 |
| **2** | 134.9, 146.3 | 145.2, 146.3 | 125.9, 127.6 | 108.9, 107.5 |
| **3** | 148.6, 156.5 | 148.6, 153.1 | 135.5, 138.9 | 132.1, 149.7 |
| **4** | 152.0, 151.4 | 149.7, 152.0 | 142.9, 142.3 | 141.7, 141.2 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Source of Variation** | **Degrees of Freedom** | **Mean Square** | **Expected Mean Square** |
| ***Technician*** | 3 | *MST* = 445 |  |
| ***Machine*** | 3 | *MSM* = 549 |  |
| ***Interaction*** | 9 | *MS*(*TM*) = 87 |  |
| ***Error*** | 16 | *MSE* = 18 |  |

1. (5pts) Find the variance components for each source of variation in this analysis. Report these in table form, with a column showing the source, a column showing the corresponding variance components, and a column expressing the variance components as percentages.

Variance Components

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Formula | Variance | Fraction of total | Percent of total |
| Error | 18 | 18 | 0.116129 | 11.61% |
| Interaction | 87 = 18 + 2 \* | 34.5 | 0.222581 | 22.26% |
| Machine | 549 = 87 + 2\*4\* | 57.75 | 0.372581 | 37.26% |
| Technician | 445 = 87 + 2\*4\* | 44.75 | 0.28871 | 28.87% |

1. (5pts) State the appropriate null hypothesis and alterative hypothesis for each effect.

Main Effect of Factor Technician:

H0:σ2technician = 0 vs. HA: σ2technician > 0

Main Effect of Factor Machine:

H0:σ2machine = 0 vs. HA: σ2machine > 0

Technician X Machine Interaction:

H0: there is no interaction    
HA: an interaction exists

1. (5pts) Conduct F tests for each of the effects in the ANOVA. (For table values of Fcritica, use the nearest value shown in Table B4 in the Appendix of the textbook.) Draw a conclusion regarding each null hypothesis.

|  |  |  |
| --- | --- | --- |
| **Source** | **F** | **F** |
| ***Technician*** | MSA / MSAB | 445 / 87 = 5.115 |
| ***Machine*** | MSB / MSAB | 549 / 87 = 6.31 |
| ***Interaction*** | MSAB / MSE | 87 / 18 = 4.83 |

d) (5pts) What test interpretation would best indicate consistency of machine performance?

The consistency of performance will be there if the variance due to main effect of machine and interaction is small. In other words if we can say that most of the variance in results can be attributed to technician – technician variance than it indicates consistency of machine performance.

If the null hypothesis H\_0: σ\_Machine^2=0 is not rejected, this would indicate a relatively high level of consistency. Predictability or consistency varies inversely with the variance.

4) In a quality control study, random samples of 4 items from each of 3 vendors chosen at random from a large population of vendors, were measured for quality.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | Df | SS | MS | EMS |
| Between Vendors | 2 | 665.2 | 332.6 |  |
| Within Vendors | 9 | 543.5 | 60.4 |  |
| Total | 11 | 1208.7 |  |  |

1. (5pts) Find the variance components for the between and within vendor sources of variation.

*To get use the method of moments:*

68.05

1. (5pts) Compute the intra-class correlation coefficient.

ICC = 68.05 / (68.05 + 60.4) = 0.5298 = 52.98%

c) (5pts) Where should they focus attention to improve consistency in quality?

A large ICC would mean that there is lot of variation from one vendor to another and a low ICC would mean that there is lot of variation within the vendor.

Since our ICC lies in the middle there has to be focus on both areas namely: improve consistency across vendors and improve process so that vendors can reduce variability within their own outputs.

*Vendors first, but there is also a lot of variance (47%) within vendors*.

5) A study measured how different planting and cooking types affected the protein content of a certain type of bean. There were two planting mechanisms and three cooking types considered. To carry out this experiment, a field was divided into six rows. For each row, a planting mechanism was randomly chosen and applied, so that each mechanism was assigned to three rows. Specifically, the planting mechanism ran down the length of the row from one end to the other; afterward it was refilled with beans and recalibrated. Afterward, three pounds of beans were harvested from each row and divided into three 1-lb batches. The cooking types were then assigned randomly to these batches. Finally, the protein content was measured for each cooked batch.

1. (10 pts) Identify (i.e., name) the treatment design and randomization design involved in this study.

Planting=fixed, cooking=fixed, row=random

Planting and cooking, factorial design. Each cooking type is combined with each planting method in the experiment.

Split-plot in a CRD. Units were assigned (randomly) a planting method. Then the row with a set planting method is split to accommodate each of the three cooking types.

Treatment design:: two-factor factorial with planting mechanism x cooking type

Randomization design: split-plot in a CRD with planting mechanism as the whole plot factor cooking type as the subplot factor

1. (5 pts) Provide the ANOVA table, including columns for source and degrees of freedom.

| **Type 3 Analysis of Variance** | | |
| --- | --- | --- |
| **Source** | **DF** |
| **planting** | 1 |
| **cooking** | 2 |
| **planting \* cooking** | 2 |
| **row(planting)** | 4 |
| **Residual** | 8 |

1. (5pts) How would the F-test for the planting mechanism be computed?

Planting F = MSplanting / MSrow(planting)