Module 9 Assignment

Alex Schlesener

2023-10-22

Contents

Overview

Consider an experimental study of generalization. A mild shock is presented to participants in the presence of a rectangle of light that is 11 inches high and 1 inch wide. Participants are then randomly divided into five experimental groups, each of which is tested in the presence of a rectangle of light, but with no shock. The independent variable is the height of the rectangle of light on the test trials; the five groups see a rectangle whose height is either 7, 9, 11, 13, or 15 inches high. An average galvanic skin response (GSR) measure is obtained for each participant. (GSR or skin conductance is a physiological measure of arousal—sweat.)

- H1: The magnitude of conditioned responses should vary directly with the magnitude of the test stimulus; in doing so, GSR scores should increase as a function of the height of the rectangle of light.
- H2: There should be a trend for GSR scores to be higher the closer the test stimulus is to the training stimulus.

Description of Data

This data is from an experimental study of generalization.The independent variable is the height of the rectangle of light on the test trials; the five groups see a rectangle whose height is either 7, 9, 11, 13, or 15 inches high. An average galvanic skin response (GSR) measure is obtained for each participant. (GSR or skin conductance is a physiological measure of arousal—sweat.)

Import Data

```
generalization.Data <- read.csv("Data/generalization.csv")

# Check the dimensions and the type of variables that we have.

str(generalization.Data)
```

```
## 'data.frame':    100 obs. of  2 variables:
##  $ gsr     : num  5.509 0.977 0.563 0.964 3.248 ...
##  $ stimulus: int  7 7 7 7 7 7 7 7 7 7 ...
```

```
summary(generalization.Data)
```

```
##       gsr              stimulus
##  Min.   : 0.06181   Min.   : 7.00
##  1st Qu.: 2.06571   1st Qu.: 9.00
##  Median : 4.07697   Median :11.00
##  Mean   : 4.02104   Mean   :11.14
##  3rd Qu.: 5.60785   3rd Qu.:13.00
##  Max.   :10.54368   Max.   :15.00
```

Convert categorical variables to factors

We must first convert the categorical variable, "stimulus", to a factor.

```
# Change Stimulus to a factor and replace the integer vector.

generalization.Data$stimulus <- factor(generalization.Data$stimulus,
                                labels=c('7in', '9in', '11in', '13in', '15in'))

# Use str() again to check the data frame.

str(generalization.Data)
```

```
## 'data.frame':    100 obs. of  2 variables:
##  $ gsr     : num  5.509 0.977 0.563 0.964 3.248 ...
##  $ stimulus: Factor w/ 5 levels "7in","9in","11in",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```
# Double check the labels that we assigned to the five levels
# of stimulus.

levels(generalization.Data$stimulus)
```

```
## [1] "7in"  "9in"  "11in" "13in" "15in"
```

```r
summary(generalization.Data)
```

```
##       gsr           stimulus
##  Min.   : 0.06181   7in :20
##  1st Qu.: 2.06571   9in :18
##  Median : 4.07697   11in:19
##  Mean   : 4.02104   13in:21
##  3rd Qu.: 5.60785   15in:22
##  Max.   :10.54368
```

Examine the data

First, we will calculate the descriptive statistics by condition and note any general trends with the means.

```r
aggregate(gsr ~ stimulus, data = generalization.Data, mean)
```

```
##   stimulus      gsr
## 1      7in 1.866905
## 2      9in 4.160242
## 3     11in 5.791177
## 4     13in 4.452991
## 5     15in 3.924382
```

```r
aggregate(gsr ~ stimulus, data = generalization.Data, sd)
```

```
##   stimulus      gsr
## 1      7in 1.717617
## 2      9in 2.367456
## 3     11in 2.101078
## 4     13in 2.165615
## 5     15in 1.808646
```

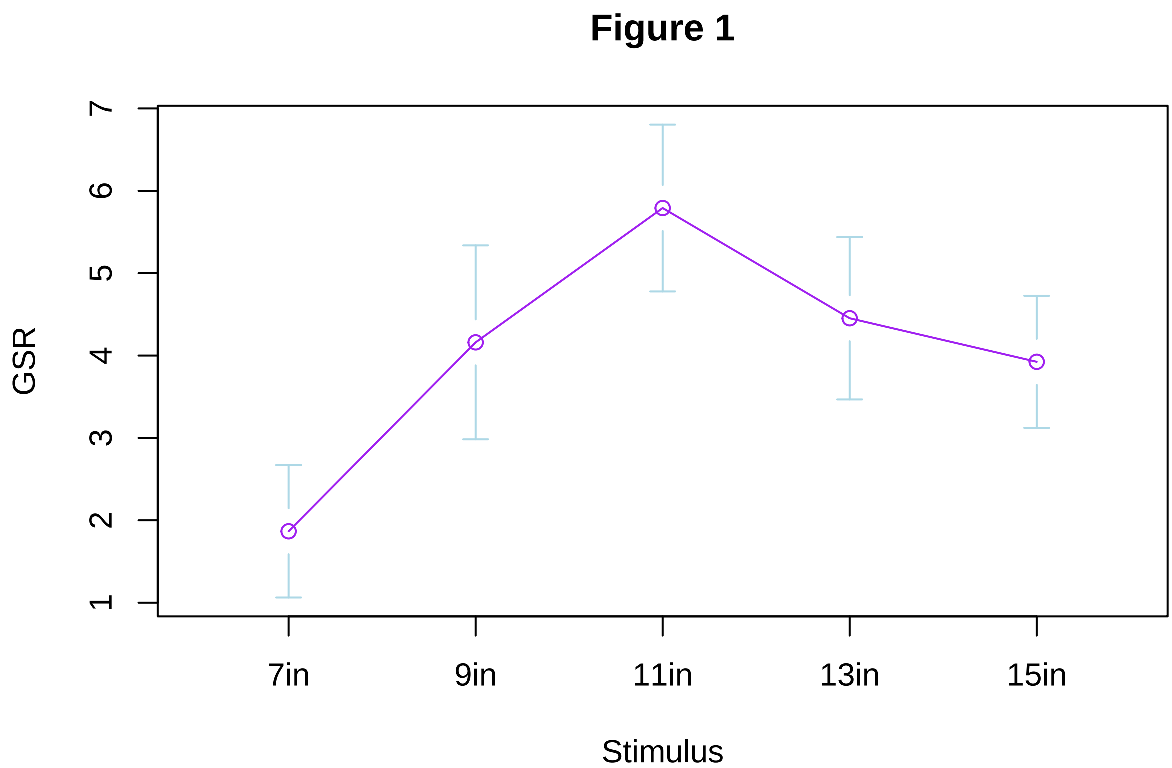
```r
aggregate(gsr ~ stimulus, data = generalization.Data, length)
```

```
##   stimulus gsr
## 1      7in  20
## 2      9in  18
## 3     11in  19
## 4     13in  21
## 5     15in  22
```

Next, we will interpret the cell means by using the 'gplots' library.

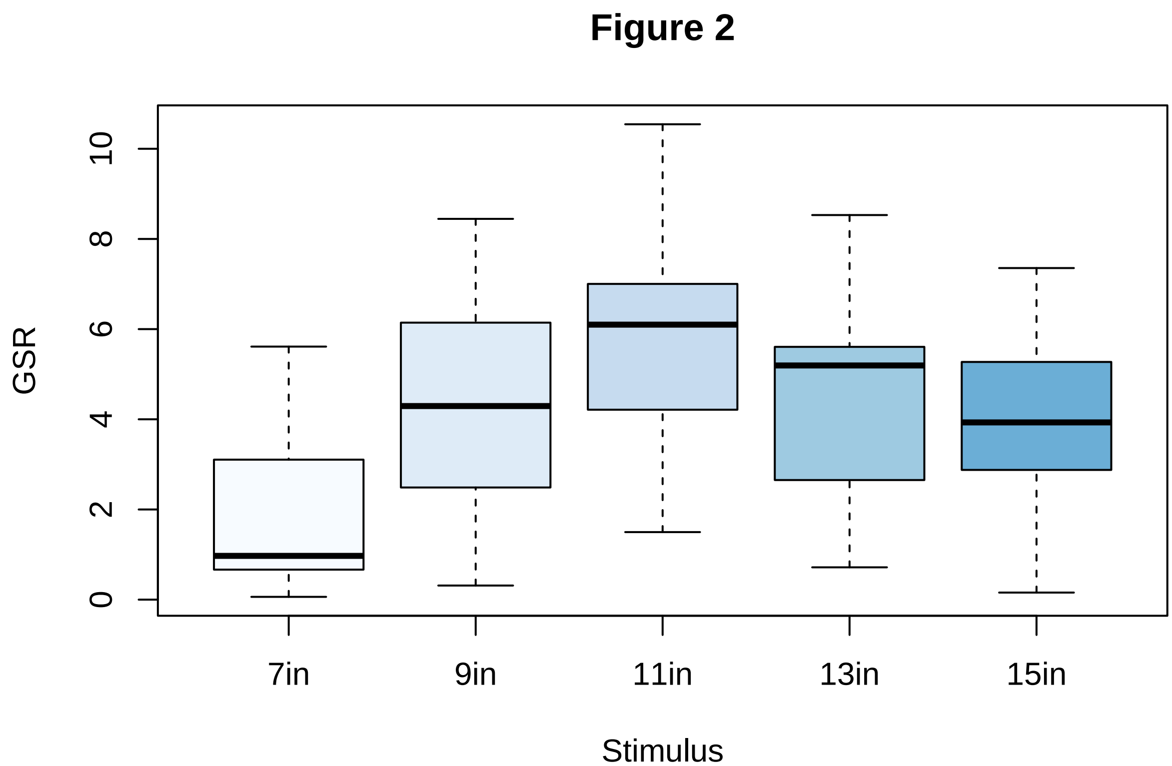
```r
library(gplots)

plotmeans(gsr ~ stimulus, generalization.Data,
          ylab="GSR", xlab="Stimulus", col="purple",
          barcol="lightblue", n.label = F, main = "Figure 1")
```



Lastly, we will examine boxplots or histograms by condition.

```
boxplot(generalization.Data$gsr ~ generalization.Data$stimulus, col=blues9,
        xlab = "Stimulus",
        ylab = "GSR",
        main = "Figure 2")
```



One-way Analysis of Variance

First, we will use lm() to conduct a one-way ANOVA.

```
fitData <- lm(gsr ~ stimulus, generalization.Data)
summary(fitData)
```

```
##
## Call:
## lm(formula = gsr ~ stimulus, data = generalization.Data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.2930 -1.2357  0.0895  1.3111  4.7525
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     1.8669     0.4550   4.103 8.60e-05 ***
## stimulus9in     2.2933     0.6611   3.469 0.000786 ***
## stimulus11in    3.9243     0.6518   6.020 3.26e-08 ***
## stimulus13in    2.5861     0.6357   4.068 9.79e-05 ***
## stimulus15in    2.0575     0.6286   3.273 0.001485 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 2.035 on 95 degrees of freedom
## Multiple R-squared:  0.2851, Adjusted R-squared:  0.255
## F-statistic:  9.47 on 4 and 95 DF,  p-value: 1.739e-06
```

```
# Obtain the ANOVA summary table
anova(fitData)
```

```
## Analysis of Variance Table
##
## Response: gsr
##           Df Sum Sq Mean Sq F value    Pr(>F)
## stimulus   4 156.81  39.203  9.4696 1.739e-06 ***
## Residuals 95 393.29   4.140
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Calculate the total sum of squares
var(generalization.Data$gsr)*(100-1)
```

```
## [1] 550.1036
```

Summary

By using summary() on the fitted model 'fitData', we see that the multiple $R\hat{}2 = 0.2851$. Therefore, in the one-way ANOVA, the *Stimulus* conditions explain about 28.51% of variance in the average galvanic skin response (GSR). Additionally, the overall F statistic was 9.47 on 4 and 95 degrees of freedom and was statistically significant (p < .001). For this reason, we reject the null hypothesis. Furthermore, the total sum-of-squares (550.10) can be separated into variability between-groups (156.81) and within-groups (393.29).

Check Homogeneity of Variance Assumption

```
library(car)
```

```
## Loading required package: carData
```

```
leveneTest(gsr ~ stimulus, data = generalization.Data, center = mean)
```

```
## Levene's Test for Homogeneity of Variance (center = mean)
##       Df F value Pr(>F)
## group  4  0.8501  0.497
##       95
```

If we are assuming that $\alpha = .05$, we can conclude that the population variance of *GSR* does not differ across the five different conditions.

Comparison among Means via Trend Analysis

```r
library(emmeans)
genStudy <- emmeans(fitData, 'stimulus')

poly.contrast <- contrast(genStudy, 'poly')
poly.contrast
```

```
##  contrast  estimate   SE df t.ratio p.value
##  linear        4.41 1.42 95   3.111  0.0025
##  quadratic    -8.61 1.70 95  -5.076  <.0001
##  cubic         1.47 1.45 95   1.015  0.3127
##  quartic       6.09 3.88 95   1.567  0.1203
```

Summary

Assuming that $\alpha = .05$, the linear and quadratic trends are statistically significant.

Effect Size

```r
# Linear Trend
linear.effect <- ((3.111)^2)/(((3.111)^2)+95)
linear.effect
```

```
## [1] 0.09245774
```

```r
# Quadratic Trend
quad.effect <- ((-5.076)^2)/(((-5.076)^2)+95)
quad.effect
```

```
## [1] 0.2133533
```

Summary

Based on the effect size, the quadratic trend of stimulus type on galvanic skin response has the largest effect.

Summary of Results

To assess whether stimulus size had an effect on galvanic skin response (GSR) measurements, we conducted a one-way analysis of variance. The dependent variable is GSR measurement and the independent variable is the height of the rectangle of light on each of the test trials (with five levels). We found overall statistical significance regarding stimulus type, $F(4,\ 95) = 9.47$, $p < .001$. In addition, the effect size ($\eta^2 = .29$) indicates about 28.51% of variance in the galvanic skin response (GSR).

Because stimulus type was found to be related to job satisfaction, we then conducted a trend analysis using orthogonal polynomials. Both the linear, $t(95) = 3.11$ ($p = .003$), and the quadratic, $t(95) = 3.11$ ($p = .003$), effects were determined to be statistically significant. However, the quadratic trend of stimulus size on galvanic skin response has a larger effect size ($\eta^2 = .21$) than the that of the linear effect ($\eta^2 = .09$). Figure 1 depicts the nature of the parabolic relationship between stimulus size and GSR measurements.