# W203-1 - Fall 2014 - Lab 3

**Juan Jose Carin**

Friday, November 21, 2014

## Contents

# Part 1: Multiple Choice

I've included comments in some sections of this Part, (not—only—to justify my choice but) to document my own reasoning when selecting the right answer, for future reference.

## 1) b) Natural experiment

Comment: *The conditions for a true experiment are not met, which discards some of the other answers. And there is a treatment condition (the photos of cats—something not selected by the experimenter), which discards an "Associatonal non-experiment."*

## 2) c) Pretest-postest experimental design

Comment: *All the conditions—the 3 treatment conditions would be the different artists, and the control condition would be the "quiet room"—are measured before and after (the treatment is applied to 3 of the 4 groups).*

## 3) f) None of the above

Comment: *Answers (a) to (c) would be true if we exchanged "parametric" by "non-parametric" and vice versa, so they are discarded, as well as (e). (d) is not always true.*

## 4) c) Dependent-sample *t*-test

Comment: *Since we are comparing the means of 2 variables measured within the same group.*

## 5) d) 42

Comment: *We know that $P(A \cap B) = P(A) \cdot P(B|A)$, and since $A$ and $B$ are independent, $P(B|A) = P(B)$, and hence $P(A \cap B) = P(A) \cdot P(B) = 0.6 \cdot 0.7 = 0.42\ldots$ if we worked with the whole population, which is not the case. For any sample (of this size: 100), the probability may differ, but its expected value will be also 0.42, so for 100 San Franciscans randomly selected, the expected value—i.e., the mean—of those who drink beer* ***and*** *wear plaid will be $\mu = n \cdot p = 100 \cdot 0.42 = 42$.*

*The following R script, where 10,000 samples are simulated, demonstrates it.*

```
## Warning: package 'knitr' was built under R version 3.1.2
```

```r
   # A simulation to prove my answer
beer <- as.data.frame(matrix(0, 100, 1e4))
plaid <- beer
beer <- apply(beer, 2, function(x){rbinom(100,1,.6)})
plaid <- apply(plaid, 2, function(x){rbinom(100,1,.7)})
both <- apply(beer*plaid, 2, sum)
mean(both)
```

```
## [1] 41.9964
```

### 6) a) Yes, because the odds of getting a type 1 error for at least one month are inflated above 5%.

Comment: *The odds of getting a type 1 error will actually be above 95%, since* $1 - .95^{\binom{12}{2}} = 1 - .95^{66} \simeq 1 - 0.0339 = 0.9661$. *A corrected t-test (using a method like Bonferroni) would be necessary.*

### 7) b) External validity

Comment: *Caused by the* ***sampling bias*** *due to a non-random sample of the population—only a specific subset of the population is represented, one with traits related to the dependent variable under study (self-confidence). Therefore, the results cannot be generalized to the whole population.*

*[Internal validity would have been threatened if we would have had* ***(sample) selection bias****, i.e., errors occurring not in the process of gathering the sample but in any process thereafter—for instance, there are differences between both groups (an unequal number of test subjects have similar subject-related variables) at pre-test, that may partially cause the observed outcome.]*

### 8) b) Difference of means, divided by the pooled standard deviation of the variable.

Comment: *I.e., Cohen's d.*

---

## Part 2: Test Selection

### 1) e) Chi-square test

Comment: *Since the outcome variable (income91) is categorical (several levels or categories), and there is a single predictor variable (visitart), which is also categorical (with 2 possible values, different entities in each category).*

### 2) d) Anova

Comment: *Since the outcome variable (age) is continuous, and the predictor variable (country) is categorical (with more than 2 possible values).*

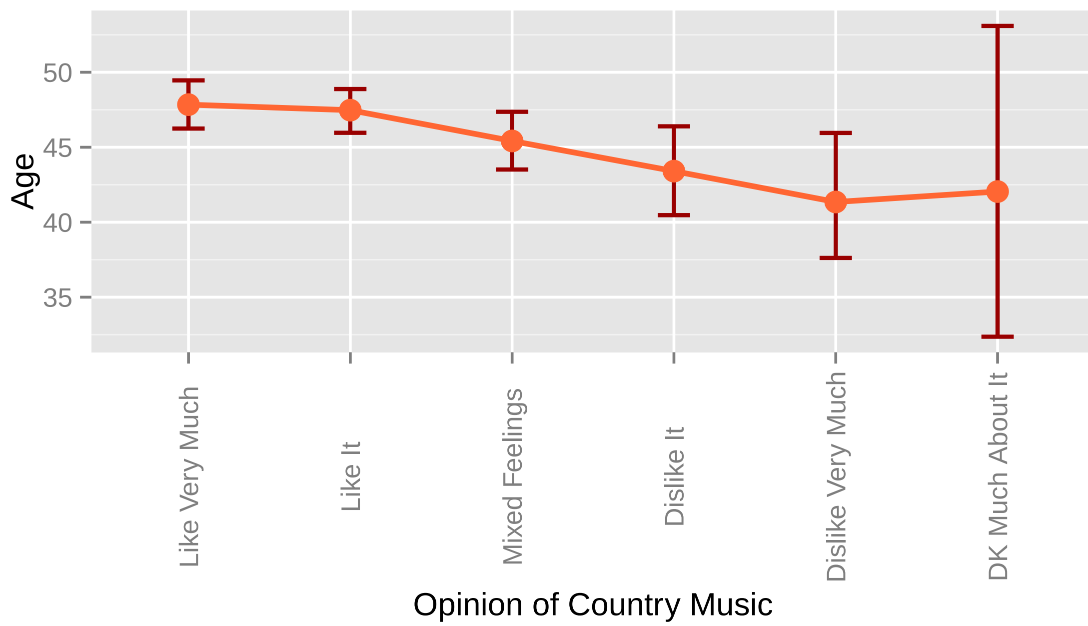


Figure 1: Error bar graph of the mean Age depending on Opinion of Country Music

## 3) a) *t*-test

Comment: *Since the outcome variable (sibs) is continuous, and the predictor variable (not relig but a new variable based on it: catholic or not) is categorical (with only 2 categories). Moreover, the assumptions of homogeneity of variance and normality are met (otherwise the Wilcoxon rank-sum test would be the answer).*

```
## sex: Male
## catholic: Catholic
##
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.9534, p-value = 0.0002099
##
## --------------------------------------------------------
## sex: Female
## catholic: Catholic
##
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.8155, p-value = 9.973e-15
##
## --------------------------------------------------------
## sex: Male
## catholic: Non-catholic
##
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.834, p-value < 2.2e-16
```

```
##
## --------------------------------------------------------
## sex: Female
## catholic: Non-catholic
##
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.8718, p-value < 2.2e-16
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   1  0.0027 0.9585
##       639
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   1  0.0103 0.9193
##       852
```

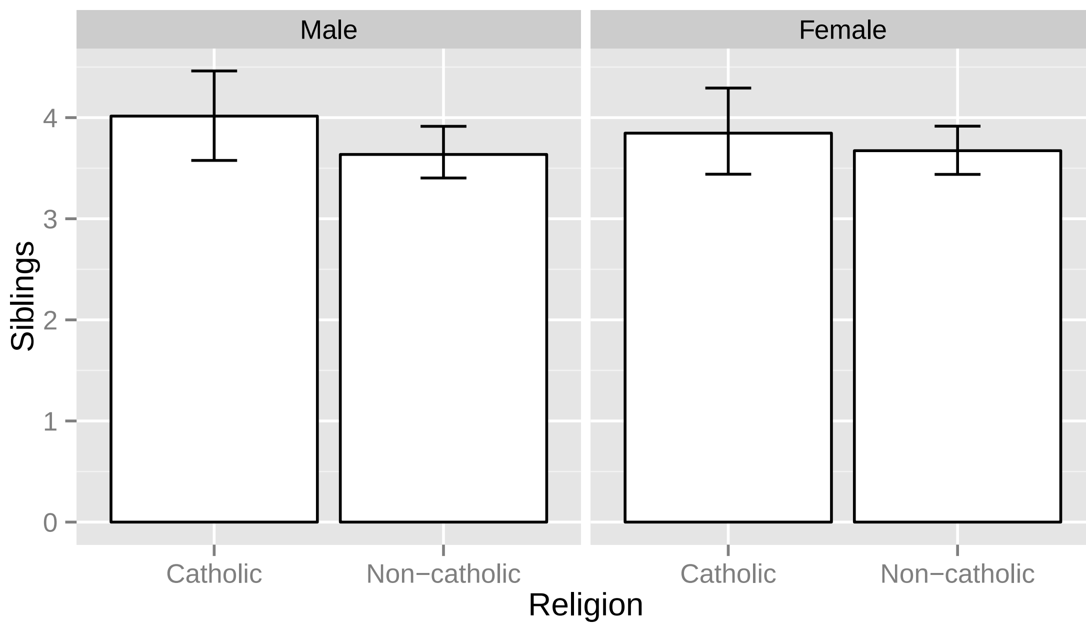


Figure 2: Bar chart of the mean number of siblings for catholic vs. non-catholic, for men and women

## 4) b) Pearson correlation

Comment: *Since both variables are continous.*

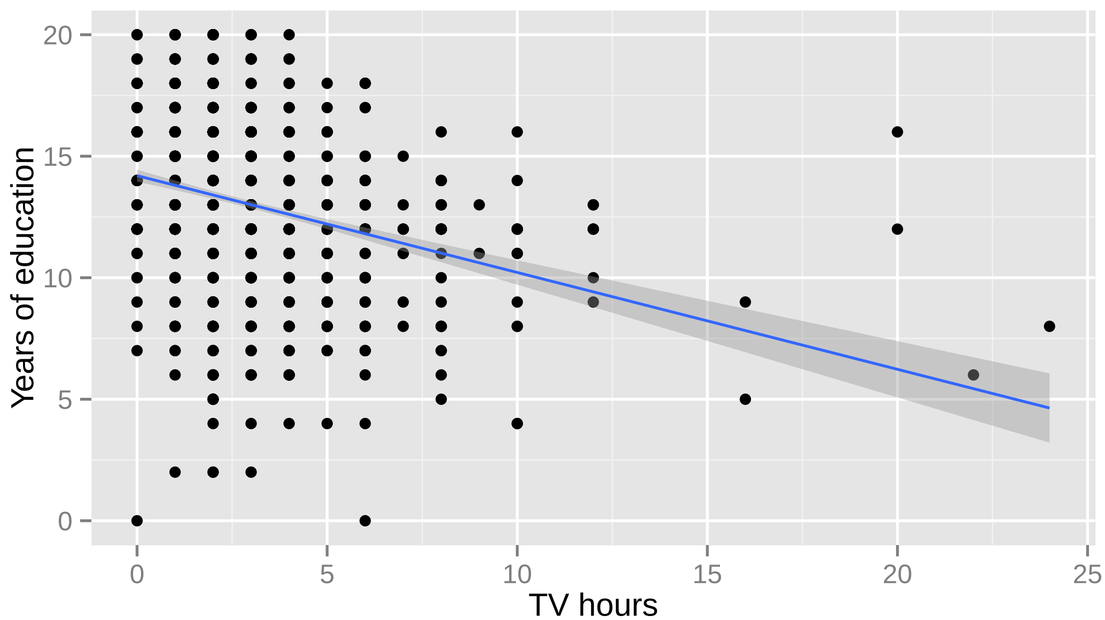


Figure 3: Scatterplot of TV hours against Years of education

### 5) d) Chi-square test

<u>Comment</u>: *Because there are 2 variables, and both are categorical.*

---

## <u>Part 3: Data Analysis and Short Answer</u>

### 1) Task 1: Chi-square test between marital status (*marital*) and political orientation (*politics*).

First of all we look for inappropriate values: *marital* has one "NA" value (so it is coded NA), and *politics* have some NA values.

```
  # Descriptive statistics, data cleansing and re-coding
summary(GSS$marital)
```

```
##       married       widowed      divorced     separated never married
##           795           165           213            40           286
##            NA
##             1
```

```r
summary(GSS$politics)
```

```
##      Liberal     Tend Lib     Moderate    Tend Cons Conservative
##          193          193          527          248          282
##         NA's
##           57
```

```r
Part3.1 <- GSS[GSS$marital != "NA" & !is.na(GSS$politics), c("marital", "politics")]
Part3.1$marital <- factor(Part3.1$marital)
summary(Part3.1)
```

```
##           marital            politics
##  married      :769   Liberal     :192
##  widowed      :149   Tend Lib    :193
##  divorced     :204   Moderate    :527
##  separated    : 39   Tend Cons   :248
##  never married:281   Conservative:282
```

After that we can run the Chi-square test (with two different commands, though the results are the same):

```r
   # Chi-square test
table(Part3.1$marital, Part3.1$politics)
```

```
##
##                  Liberal Tend Lib Moderate Tend Cons Conservative
##   married             93       92      271       140          173
##   widowed             15       16       57        24           37
##   divorced            22       36       79        38           29
##   separated            7        3       22         6            1
##   never married       55       46       98        40           42
```

```r
(cs = chisq.test(Part3.1$marital, Part3.1$politics))
```

```
##
##  Pearson's Chi-squared test
##
## data:  Part3.1$marital and Part3.1$politics
## X-squared = 44.2255, df = 16, p-value = 0.0001823
```

```r
CrossTable(Part3.1$marital, Part3.1$politics, fisher = FALSE, chisq = TRUE,
           expected = TRUE, asresid = TRUE, prop.t = FALSE, format = "SPSS")
```

```
##
##    Cell Contents
## |-------------------------|
## |                   Count |
## |         Expected Values |
## | Chi-square contribution |
## |             Row Percent |
## |          Column Percent |
## |            Adj Std Resid |
## |-------------------------|
##
## Total Observations in Table:  1442
```

```
##
##                  | Part3.1$politics
## Part3.1$marital |      Liberal  |     Tend Lib  |     Moderate  |    Tend Cons  | Conservative  |    Row Total |
## ----------------|--------------|--------------|--------------|--------------|--------------|--------------|
##         married |           93  |           92  |          271  |          140  |          173  |          769  |
##                 |      102.391  |      102.924  |      281.042  |      132.255  |      150.387  |              |
##                 |        0.861  |        1.160  |        0.359  |        0.454  |        3.400  |              |
##                 |      12.094% |      11.964% |      35.241% |      18.205% |      22.497% |      53.329% |
##                 |      48.438% |      47.668% |      51.423% |      56.452% |      61.348% |              |
##                 |       -1.459  |       -1.694  |       -1.101  |        1.083  |        3.009  |              |
## ----------------|--------------|--------------|--------------|--------------|--------------|--------------|
##         widowed |           15  |           16  |           57  |           24  |           37  |          149  |
##                 |       19.839  |       19.942  |       54.454  |       25.626  |       29.139  |              |
##                 |        1.180  |        0.779  |        0.119  |        0.103  |        2.121  |              |
##                 |      10.067% |      10.738% |      38.255% |      16.107% |      24.832% |      10.333% |
##                 |       7.812% |       8.290% |      10.816% |       9.677% |      13.121% |              |
##                 |       -1.232  |       -1.002  |        0.457  |       -0.373  |        1.715  |              |
## ----------------|--------------|--------------|--------------|--------------|--------------|--------------|
##        divorced |           22  |           36  |           79  |           38  |           29  |          204  |
##                 |       27.162  |       27.304  |       74.555  |       35.085  |       39.895  |              |
##                 |        0.981  |        2.770  |        0.265  |        0.242  |        2.975  |              |
##                 |      10.784% |      17.647% |      38.725% |      18.627% |      14.216% |      14.147% |
##                 |      11.458% |      18.653% |      14.991% |      15.323% |      10.284% |              |
##                 |       -1.148  |        1.930  |        0.698  |        0.584  |       -2.076  |              |
## ----------------|--------------|--------------|--------------|--------------|--------------|--------------|
##       separated |            7  |            3  |           22  |            6  |            1  |           39  |
##                 |        5.193  |        5.220  |       14.253  |        6.707  |        7.627  |              |
##                 |        0.629  |        0.944  |        4.211  |        0.075  |        5.758  |              |
##                 |      17.949% |       7.692% |      56.410% |      15.385% |       2.564% |       2.705% |
##                 |       3.646% |       1.554% |       4.175% |       2.419% |       0.355% |              |
##                 |        0.864  |       -1.058  |        2.612  |       -0.304  |       -2.712  |              |
## ----------------|--------------|--------------|--------------|--------------|--------------|--------------|
##   never married |           55  |           46  |           98  |           40  |           42  |          281  |
##                 |       37.415  |       37.610  |      102.696  |       48.327  |       54.953  |              |
##                 |        8.265  |        1.872  |        0.215  |        1.435  |        3.053  |              |
##                 |      19.573% |      16.370% |      34.875% |      14.235% |      14.947% |      19.487% |
##                 |      28.646% |      23.834% |      18.596% |      16.129% |      14.894% |              |
##                 |        3.441  |        1.638  |       -0.648  |       -1.467  |       -2.171  |              |
## ----------------|--------------|--------------|--------------|--------------|--------------|--------------|
##    Column Total |          192  |          193  |          527  |          248  |          282  |         1442  |
##                 |      13.315% |      13.384% |      36.546% |      17.198% |      19.556% |              |
## ----------------|--------------|--------------|--------------|--------------|--------------|--------------|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## ------------------------------------------------------------
## Chi^2 =  44.2255     d.f. =  16     p =  0.0001822704
##
##
##
##        Minimum expected frequency: 5.192788
```

### A. What are the null and alternative hypothesis for your test?

The **null-hypothesis** is that there is no relationship between the two categorical variables (i.e., between any of their categories), that have been cross-tabulated. In that case there would be no significant difference between the expected frequencies (those expected by chance alone) and the observed frequencies in any category. In other words, it means that the observed cell counts are independent from one another.

The **alternative hypothesis** is the opposite: a relationship exists between any of the categories, so there is a significant difference between the expected and the observed frequencias at least in one case, etc. So in this particular case the alternative hypothesis would be true if there were an association between marital status and political orientation (whatever it might be: e.g., a significant number of conservative people are married, or those individuals that are liberal tend to never get married more than other people. . . )

### B. What test statistic and p-value do you get?

As shown before (at the top of this page and also in page 8), the **Chi-square test statistic** is **44.2254972** and the **p-value** is **1.8e-04** (highly statistically significant).

```
cs
```

```
##
##  Pearson's Chi-squared test
##
## data:  Part3.1$marital and Part3.1$politics
## X-squared = 44.2255, df = 16, p-value = 0.0001823
```

That means that there are significant differences between the observed and expected values (and hence a relationship exists between political orientation and marital status).

We can rely on the results of this test, since **the two important assumptions are met**: each person contributes to only cell of the contingency table (categories are mutually exclusive and exhaustive), and the expected frequencies are greater than 5 in all cases (the minimum expected frequency is 5.193, as can be seen at the top of this page).

```
cs$expected
```

```
##                Part3.1$politics
## Part3.1$marital    Liberal   Tend Lib  Moderate  Tend Cons Conservative
##   married       102.391123 102.924411 281.04230 132.255201   150.386963
##   widowed        19.839112  19.942441  54.45423  25.625520    29.138696
##   divorced       27.162275  27.303745  74.55479  35.084605    39.894591
##   separated       5.192788   5.219834  14.25312   6.707351     7.626907
##   never married  37.414702  37.609570 102.69556  48.327323    54.952843
```

### C. Conduct an effect size calculation for your relationship.

Our contingency table is much larger than 2x2, so Odd Ratios would not be so appropriate. We'll use **Cramer's V** instead.

```
## [1] Cramer's V: 0.0876
```

Since its value is so small (under 0.2), we can say that **the strength of association between the two variables is weak**.

**D. Evaluate your hypothesis in light of your tests of statistical and practical significance. What, if anything, can you conclude from your results?**

There is a statistically significant association between political orientation and marital status: $\chi^2(16) = 44.23, p < .001$. Hence, we are able to reject the null hypothesis. However, the practical significance of this finding is very small: $V = 0.088$, so the observed effect is unimportant, not meaningful, and not of real interest.

## 2) Task 2: Pearson correlation analysis between age when married (*agewed*) and hours of tv watched (tvhours).

First of all we look for inappropriate values, and code them as NA: 98 and 99, and also 0 for *agewed*.

```
   # Descriptive statistics, data cleansing and re-coding
GSS$agewed[GSS$agewed == 0 | GSS$agewed == 98 | GSS$agewed == 99] <- NA
GSS$tvhours[GSS$tvhours == 98 | GSS$tvhours == 99] <- NA
Part3.2 <- GSS[!is.na(GSS$agewed) & !is.na(GSS$tvhours), c("agewed", "tvhours")]
summary(Part3.2)
```

```
##      agewed         tvhours
##  Min.   :13.00   Min.   : 0.000
##  1st Qu.:19.00   1st Qu.: 2.000
##  Median :22.00   Median : 2.000
##  Mean   :22.77   Mean   : 2.902
##  3rd Qu.:25.00   3rd Qu.: 4.000
##  Max.   :58.00   Max.   :24.000
```

After that we check the assumption of normality (and plot both variables):

```
##
##  Shapiro-Wilk normality test
##
## data:  Part3.2$agewed
## W = 0.8882, p-value < 2.2e-16
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Part3.2$tvhours
## W = 0.7776, p-value < 2.2e-16
```

Both Shapiro tests are highly significant, so the 2 variables meet the assumption of normality. From the scatterplot we can deduce that the correlation will be very small (and negative—the later someone gets the married, the less hours he or she watches TV… but the relationship between both variables is very small, most of the people get married between 15 and 40, and watches TV between 0 and 10 hours, forming a round cluster.

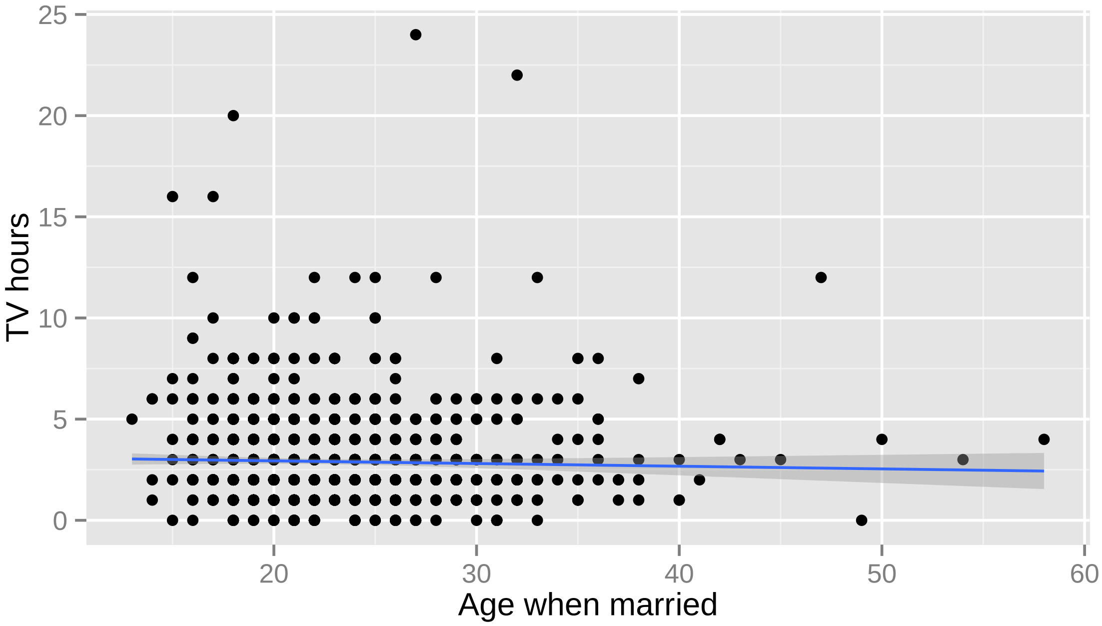


Figure 4: Scatterplot of Age when married against TV hours

Now we can run the Pearson correlation analysis:

```r
    # Pearson correlation
(correlation <- cor.test(Part3.2$agewed, Part3.2$tvhours))
```

```
##
##  Pearson's product-moment correlation
##
## data:  Part3.2$agewed and Part3.2$tvhours
## t = -1.0349, df = 1192, p-value = 0.3009
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.08654554  0.02681630
## sample estimates:
##         cor
## -0.02996096
```

#### A. What are the null and alternative hypothesis for your test?

The **null-hypothesis** is that the correlation coefficient is equal to zero, i.e., there is **not** a linear relationship between the two variables. The **alternative hypothesis** is the opposite: a linear relationship exists, so any change in one variable involves (but does not necessarily cause!!) a proportional change in the other variable, in the same or the opposite direction.

#### B. What test statistic and p-value do you get?

As shown in the previous page, the **Pearson correlation coefficient test statistic** is **-1.0348774** (relatively small in absolute value—for a sample of this size, the absolute value should be close to zero to obtain a statistically significant result), and the **p-value** is **0.3009361** (not statistically significant at all).

```r
correlation$statistic
```

```
##         t
## -1.034877
```

```r
correlation$p.value
```

```
## [1] 0.3009361
```

**C. Evaluate your hypothesis in light of your tests of statistical and practical significance. What, if anything, can you conclude from your results?**

There is not a significant relationship between the age at which people get married (for the first time) and the number of hours they watch TV, $r = -0.03, p = 0.3$. Therefore, the **practical significance is null** ($r$ should have to be ten times greater to have just a small practical significance), **as well as the statistical significance** (the $p$-value is much greater than 0.05). For that reason, **we cannot reject the null hypothesis**: both variables are uncorrelated.

## 3) Task 3: Wilcox rank-sum test to determine whether a new dummy variable (*married*) is associated with the number of children (*childs*) for respondents who are 23 years old.

```r
summary(GSS$marital)
```

```
##       married       widowed      divorced     separated never married 
##           795           165           213            40           286 
##            NA 
##             1
```

```r
GSS$married <- factor(ifelse(GSS$marital == "married", 1, 0))
(Part3.3 <- subset(GSS, age == 23, select = c("childs", "married")))
```

```
##      childs married
## 12        0       0
## 29        0       0
## 65        0       0
## 68        0       0
## 326       0       0
## 358       0       1
## 415       1       1
## 459       0       0
## 495       0       0
## 690       1       1
## 700       1       0
## 919       2       1
## 951       0       0
## 972       0       0
## 973       0       0
```

```
## 994         1         1
## 998         0         0
## 1034        0         0
## 1075        1         1
## 1083        1         0
## 1206        1         1
## 1274        1         0
## 1286        0         0
## 1446        0         0
## 1447        0         0
## 1468        0         0
## 1475        0         0
## 1488        3         1
```

**A. What is the mean of your new "married" variable among 23-year-olds (e.g.,the proportion of cases in the category coded "1")?**

```
(married23_mean <- sum(Part3.3$married==1)/length(Part3.3$married))
```

```
## [1] 0.2857143
```

28.57% of the 23-years-olds (8 out of 28).

**B. What is the null and alternative hypotheses for your test?**

The **null-hypothesis** is that both groups (23-years-olds currently married or not) are equal in terms of the predictor variable (the number of children they have), i.e., the probability distributions of both groups are equal: there is a symmetry between groups with respect to probability of random drawing of a larger observation.

The **alternative hypothesis** is that a particular group have larger values of the predictor variable than the other: the probability of an observation from one group exceeding an observation from the second group (after exclusion of ties) is not equal to 0.5. In this case, that 23-years-olds are more likely to have more children, or vice versa.

**C. What test statistic and p-value do you get?**

```
table(Part3.3$childs, Part3.3$married)
```

```
##
##       0  1
##    0 17  1
##    1  3  5
##    2  0  1
##    3  0  1
```

```
(w <- wilcox.test(childs ~ married, Part3.3))
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  childs by married
## W = 19, p-value = 0.0002656
## alternative hypothesis: true location shift is not equal to 0
```

As shown above, the **test statistic** is **19**, and the **p-value** is **2.7e-04** (highly statistically significant).

### D. Conduct an effect size calculation for your relationship.

The effect size can be calculated as $r = \frac{z}{\sqrt{N}}$, where $z$ is the $z$-statistic associated with the p-value that was obtained, and $N$ is the total number of observations (in both groups). After creating a simple function that calculates this value, we obtain:

```
## childs by married Effect size, r =  -0.6891632
```

Since its absolute value is greater than 0.5, we can be sure that the effect size is large.

### E. Evaluate your hypothesis in light of your tests of statistical and practical significance. What, if anything, can you conclude from your results?

As mentioned, the effect size is large, and the p-value was below 0.001, so the Wilcox sum-rank test is both highly statisticaly and practically significant. Wen can reject the null hypothesis, and say that the number of children of currently married 23-years-olds ($Mdn$ = *1*) differ significantly from the number of children of people the same age that are not currently married ($Mdn$ = *0*), $W$ = *19*, $p < 0.001$, $r$ = *-0.69*.

## 4) Task 4: Analysis of variance between religious affiliation (*relig*) and age when married (*agewed*).

We had already properly coded *agewed* so we just need to check *relig*:

```r
summary(GSS$relig)
```

```
## Protestant   Catholic     Jewish       None      Other         DK
##        953        333         31        140         35          1
##         NA
##          7
```

```r
GSS$relig[GSS$relig == "NA" | GSS$relig == "DK"] <- NA
GSS$relig <- factor(GSS$relig)
Part3.4 <- GSS[!is.na(GSS$relig) & !is.na(GSS$agewed), c("relig", "agewed")]
```

The next step, after performing the analysis of variance, is to check the assumptions under the $F$-statistic is realiable: normality of distributions within groups, and homogeneity of variance:

```
## Part3.4$relig: Protestant
##
##  Shapiro-Wilk normality test
```

```
##
## data:  dd[x, ]
## W = 0.8522, p-value < 2.2e-16
##
## --------------------------------------------------------
## Part3.4$relig: Catholic
##
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.9175, p-value = 6.316e-11
##
## --------------------------------------------------------
## Part3.4$relig: Jewish
##
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.935, p-value = 0.1402
##
## --------------------------------------------------------
## Part3.4$relig: None
##
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.9458, p-value = 0.0006409
##
## --------------------------------------------------------
## Part3.4$relig: Other
##
##  Shapiro-Wilk normality test
##
## data:  dd[x, ]
## W = 0.9631, p-value = 0.4571
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##         Df F value Pr(>F)
## group    4  0.8521 0.4922
##       1191
```

There are 2 groups ("Jewish" and "Other") for which the assumption of normality is broken, but in the case of ANOVA that is not so serious, because the $F$-statistic controls the type 1 error well under the condition of non-formality. The really important assumption is the homogeneity of variance, and that one is met (the result is not significant at all, $p$-value $= 0.49$), so we can perform ANOVA.

```
##               Df Sum Sq Mean Sq F value   Pr(>F)
## relig          4    811  202.64   8.197 1.61e-06 ***
## Residuals   1191  29444   24.72
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

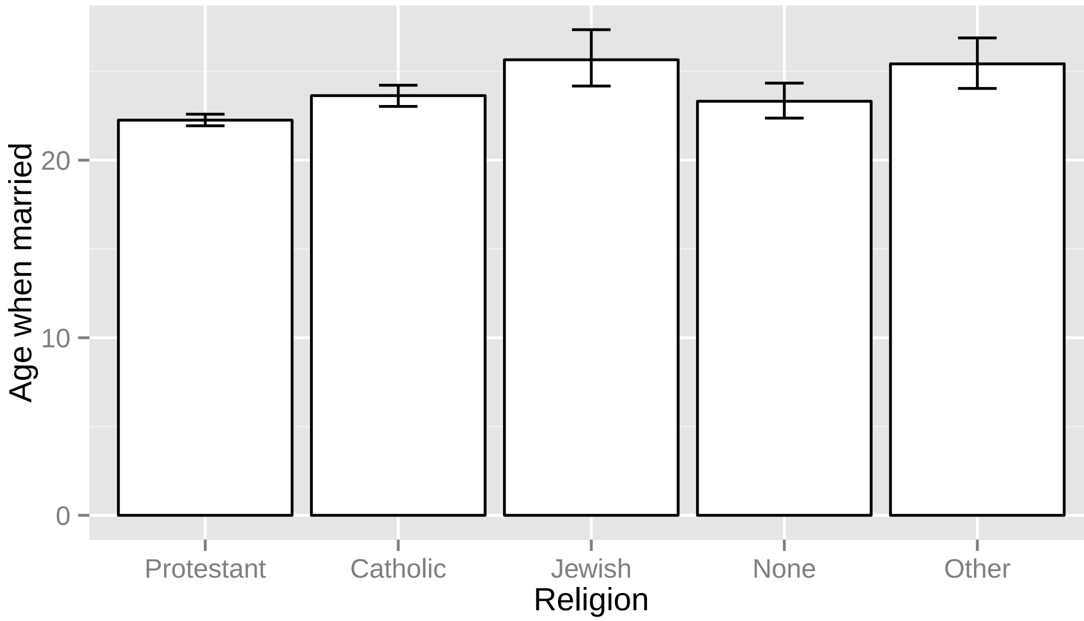


Figure 5: Bar chart of the mean age when married for different religions

### A. What is the null and alternative hypotheses for your test?

Since no specific contrasts are given, the **null hypothesis** is that all groups are simply random samples of the same population, and hence the mean of the continuous variable (*agewed*) does not differ per group (*relig*). The **alternative hypothesis** is that the religion has a statistically significant effect on the age at which people get married, and hence the means per group differ.

### B. What test statistic and p-value do you get?

As shown above, the ***F*-statistic** is **8.1967903**, and the ***p*-value** is **1.6e-06** (highly statistically significant).

### C. Are there statistically significant differences between individual pairs of groups, and if so, how do you know?

Since there are no specific hypothesis, we have to perform a **post-hoc test** or **pairwise comparison**, i.e., a series of $t$-tests between all of the different groups to search for specific differences between groups.

In order to control the familywise error rate, some correction must be applied. Some of them are tested below:

First, the classical Bonferroni correction, which is the most conservative to guarantee a low Type 1 Error (by losing Statistical Power).

```
##          Protestant Catholic Jewish  None
## Catholic 0.001 ***
## Jewish   0.013 *    0.621
## None     0.493      1        0.434
## Other    0.014 *    0.802    1       0.558
```

The *sig_stars* function shown in the Asynchronous Material has been used, so *** indicates a $p$-value below 0.001, ** a $p$-value below 0.01, and * a $p$-value below 0.05. The *pairwise.t.test* function already corrects the $p$-values (multiplying them by the corresponding correction factor), so we search for at leat one *, as usual.

The Holm method, which is not so conservative, is also tested:

```
##          Protestant Catholic Jewish  None
## Catholic 0.001 ***
## Jewish   0.011 *    0.304
## None     0.304      1        0.304
## Other    0.011 *    0.304    1       0.304
```

And so is the Benjamini-Hochberg method, which focuses on minimizing the **FDR**:

```
##      Protestant     Catholic      Jewish        None
##   0.001 ***:1            :1           :2           :3
##   0.005 ** :2     0.089 .:1    0.089 .:1    0.089 .:1
##   0.089 .  :1     0.1    :1    0.872  :1
##                   0.659  :1
```

Tukey's HSD method is not appropriate since sample sizes are very dissimilar.

As we can see, all of them give similar results: there are significant differences between 3 pairs of groups: Protestant against Catholic—the highest statistically significant result according to all the methods—, Jewish and Other. This complies with the previous bar chart, where we can see these are the only cases in which the SE bars do not overlap.

**D. Evaluate your hypothesis in light of your tests of statistical and practical significance. What, if anything, can you conclude from your results?**

First of all, we need to estimate the practical significance. A function that calculates both $R^2$ and $\omega^2$ is defined, with which we obtain the following results:

```
   # Definition and use of omega-squared function (also calculates R-squared)
      # to estimate the effect size of ANOVA analysis
omega <- function(aov.summary){
   dfm <- aov.summary[[1]][1,1]
   SSm <- aov.summary[[1]][1,2]
   SSr <- aov.summary[[1]][2,2]
   SSt <- SSm + SSr
   MSr <- aov.summary[[1]][2,3]
   omega2 <- (SSm-(dfm*MSr))/(SSt+MSr)
   print.noquote(paste("Omega-squared:", round(as.numeric(omega2), 2)))
   R2 <- SSm/SSt
   print.noquote(paste("R-squared:", round(as.numeric(R2), 2)))
}

omega(summary(aovm))
```

```
## [1] Omega-squared: 0.02
## [1] R-squared: 0.03
```

$R^2$ is close to zero, and $\omega^2$ (which is more accurate since $R^2$ is slightly biased) lays between 0.01 and 0.06, so the practical significance is small.

But we are more interested in estimating the effect size not of the overall ANOVA, but for the differences between pairs of groups. So the *mes* function of the package *computes.es* was also used to calculate Cohen's *d*.

```
## Mean Differences ES:
##
##  d [ 95 %CI] = -0.28 [ -0.42 , -0.14 ]
##   var(d) = 0.01
##   p-value(d) = 0
##   U3(d) = 39.14 %
##   CLES(d) = 42.27 %
##   Cliff's Delta = -0.15
##
##  g [ 95 %CI] = -0.28 [ -0.42 , -0.14 ]
##   var(g) = 0.01
##   p-value(g) = 0
##   U3(g) = 39.15 %
##   CLES(g) = 42.28 %
##
##  Correlation ES:
##
##  r [ 95 %CI] = -0.12 [ -0.18 , -0.06 ]
##   var(r) = 0
##   p-value(r) = 0
##
##  z [ 95 %CI] = -0.12 [ -0.18 , -0.06 ]
##   var(z) = 0
##   p-value(z) = 0
##
##  Odds Ratio ES:
##
##  OR [ 95 %CI] = 0.61 [ 0.47 , 0.78 ]
##   p-value(OR) = 0
##
##  Log OR [ 95 %CI] = -0.5 [ -0.75 , -0.25 ]
##   var(lOR) = 0.02
##   p-value(Log OR) = 0
##
##  Other:
##
##  NNT = -14.7
##  Total N = 1052
```

```
## Mean Differences ES:
##
##  d [ 95 %CI] = -0.68 [ -1.1 , -0.26 ]
##   var(d) = 0.05
##   p-value(d) = 0
##   U3(d) = 24.8 %
##   CLES(d) = 31.51 %
##   Cliff's Delta = -0.37
##
##  g [ 95 %CI] = -0.68 [ -1.1 , -0.26 ]
##   var(g) = 0.04
##   p-value(g) = 0
##   U3(g) = 24.82 %
##   CLES(g) = 31.52 %
```

```
##
##  Correlation ES:
##
##  r [ 95 %CI] = -0.11 [ -0.18 , -0.04 ]
##   var(r) = 0
##   p-value(r) = 0
##
##  z [ 95 %CI] = -0.11 [ -0.18 , -0.04 ]
##   var(z) = 0
##   p-value(z) = 0
##
##  Odds Ratio ES:
##
##  OR [ 95 %CI] = 0.29 [ 0.14 , 0.62 ]
##   p-value(OR) = 0
##
##  Log OR [ 95 %CI] = -1.24 [ -1.99 , -0.48 ]
##   var(lOR) = 0.15
##   p-value(Log OR) = 0
##
##  Other:
##
##  NNT = -7.35
##  Total N = 810

## Mean Differences ES:
##
##  d [ 95 %CI] = -0.87 [ -1.26 , -0.47 ]
##   var(d) = 0.04
##   p-value(d) = 0
##   U3(d) = 19.3 %
##   CLES(d) = 26.99 %
##   Cliff's Delta = -0.46
##
##  g [ 95 %CI] = -0.87 [ -1.26 , -0.47 ]
##   var(g) = 0.04
##   p-value(g) = 0
##   U3(g) = 19.32 %
##   CLES(g) = 27.01 %
##
##  Correlation ES:
##
##  r [ 95 %CI] = -0.15 [ -0.22 , -0.08 ]
##   var(r) = 0
##   p-value(r) = 0
##
##  z [ 95 %CI] = -0.15 [ -0.22 , -0.08 ]
##   var(z) = 0
##   p-value(z) = 0
##
##  Odds Ratio ES:
##
##  OR [ 95 %CI] = 0.21 [ 0.1 , 0.42 ]
##   p-value(OR) = 0
##
##  Log OR [ 95 %CI] = -1.57 [ -2.29 , -0.86 ]
##   var(lOR) = 0.13
##   p-value(Log OR) = 0
##
```

```
## Other:
##
## NNT = -6.4
## Total N = 813
```

The **ANOVA analysis** revealed a highly statistically significant effect of the religion on the age at which people get married, $F(4, 1191) = 8.197, p < 0.00001$. However, the practical significance is very small, $\omega^2 = 0.02$.

The **post-hoc tests** revealed that there are statistically significant differences between Protestants, and Catholics ($p < 0.001, d = -0.28$), Jews ($p < 0.05, d = -0.68$) and Others ($p < 0.05, d = -0.87$). I.e., the effect are medium-large between Protestants and Jewish and between Protestants and Others, and small-medium between Protestants and Catholics.