Psychology 205: Research Methods in Psychology
Statistics Problem Set 1 answers
and a review of basic statistics

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

April, 2021

Outline

Statistics and model fitting

1. The basic concept is that Data = Model + Residual
2. We want to find that model of the data that minimizes the Residual
3. The mean is a model of the data: It is the estimate that minimizes the residual (and the squared residual)
4. If our data are **X** then the arithmetic mean $\bar{x} = \frac{\Sigma(X_i)}{N}$ and the residual $= x_i = X_i - \bar{x}$ will have a mean of 0.
5. These residuals are also known as deviation scores (x_i).
6. We call the average squared residual the Variance, $\sigma^2 = \frac{\Sigma(X_i - \bar{x})}{N-1} = \frac{\Sigma x_i^2}{N-1}$. We divide by N-1 rather than N to get an unbiased estimate of the variance.
7. The standard deviation just the square root of the variance and is thus the Root Mean Square deviation $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma(X_i - \bar{x})^2}{N-1}} = \sqrt{\frac{\Sigma x_i^2}{N-1}}$

Finding the mean, variance and sd in R

R code

```
mean(X)
var(X)
sd(X)
describe(X) will return mean, sd, as well as median, range, min, max,
        standard error, etc.
 #for example
placebo=c(24,25,27,26,26,22,21,22,23,25,25,25)   #use c to concatenate
sum(placebo)
sum(placebo)/length(placebo)
mean(placebo)
var(placebo)
sd(placebo)
describe(placebo)
```

```
placebo=c(24,25,27,26,26,22,21,22,23,25,25,25)
 sum(placebo)
[1] 291
> sum(placebo)/length(placebo)
[1] 24.25
> mean(placebo)
[1] 24.25
> var(placebo)
[1] 3.477273
> sd(placebo)
[1] 1.864745
> describe(placebo)
   vars  n  mean   sd median trimmed  mad min max range  skew kurtosis   se
X1    1 12 24.25 1.86     25    24.3 1.48  21  27     6 -0.33    -1.33 0.54
```

Problem 1

An investigator believes that caffeine facilitates performance on a simple spelling test. Two groups of subjects are given either 200 mg of caffeine or a placebo. What test should be applied to see if these two groups differ if the results are as seen in Table 1:

Table: The effect of caffeine on spelling performance

placebo	caffeine
24	24
25	29
27	26
26	23
26	25
22	28
21	27
22	24
23	27
25	28
25	27
25	26

Problem 1 is a comparison of two groups. The data are numeric and so we want to compare if their means differ. Then we want to test how large this difference is compared to chance. We will do a t-test. We do this analysis by using R.

Comparing two groups

1. When observing data from two (or more) groups, the question is not just whether the means differ, but do they differ more than we would expect by chance if we had taken two samples from the same population.
2. That is, do our samples represent samples from different populations or the same population?
3. In our case, does the administration of caffeine change spelling ability?
4. Anytime we take multiple samples from the same population, we expect them to differ (somewhat). But by how much?
5. The expected standard error of a sample mean is $se = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
6. If we think of the placebo condition as our population, we can take many samples from the placebo condition and plot the distribution of these sample means.
7. We could then compare this to the distribution of the caffeine condition

Problem 1 using R

Problem 1 is a comparison of two groups. The data are numeric and so we want to compare if their means differ. Then we want to test how large this difference is compared to chance. We will do a t-test. We do this analysis by using R.

First, I show the R code, and then I show the results.

R code

```
#first create a data frame to hold the data.
#We use the concatenate operator (c)
#We separate the numbers with commas.

prob1 <- data.frame(placebo=c(24,25,27,26,26,22,21,22,23,25,25,25),
     caffeine =c(24,29,26,23,25,28,27,24,27,28,27,26))

prob1 #this shows us the data
dim(prob1) #how many rows and columns?
describe(prob1)  #descriptive statistics
```

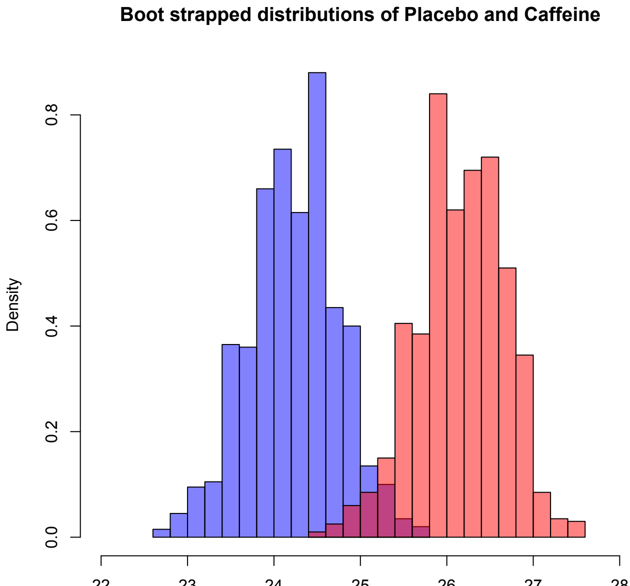
```
dim(prob1)
[1] 12  2
describe(prob1)  #descriptive statistics
         vars  n  mean   sd median trimmed  mad min max range  skew kurtosis   se
placebo     1 12 24.25 1.86   25.0    24.3 1.48  21  27     6 -0.33    -1.33 0.54
caffeine    2 12 26.17 1.85   26.5    26.2 2.22  23  29     6 -0.22    -1.33 0.53
```

Boot strap resampling

R code

```
#do some sampling to show what we expect

set.seed(42) #to get the same "random results"
boot.placebo <- matrix(sample(prob1$placebo,12000,replace=TRUE),nrow=12)
placebo.means  <- colMeans(boot.placebo)
describe(placebo.means)  #note that the standard deviation of these means is ro
   vars    n  mean   sd median trimmed  mad   min   max range  skew kurtosis
X1    1 1000 24.24 0.52  24.25   24.24 0.49 22.67 25.75  3.08 -0.09    -0.05 0

#set up a transparent color
bluecol <- rgb(0, 0, 255, max = 255, alpha = 125, names = "blue50")
#Draw the histogram
hist(placebo.means,breaks=21,xlim=c(22,28),xlab="Spelling score",
         main="Boot strapped distributions of Placebo and Caffeine",col=bluecol

#now generate the samples from the caffeine "population"
boot.caffeine <- matrix(sample(prob1$caffeine,12000,replace=TRUE),nrow=12)
caffeine.means <- colMeans(boot.caffeine)
describe(caffeine.means)

#another transparent color
redcol <- rgb(255, 0, 0, max = 255, alpha = 125, names = "red50")
hist(caffeine.means,add=TRUE,col=redcol, breaks=21,freq=FALSE )
```

1000 bootstrapped samples from each group



Gossett did not have R but worked out the t-test

R code

```
#we specify that we have equal variances
with(prob1, t.test(placebo,caffeine, equal.var=TRUE))
```

```
        Welch Two Sample t-test

data:  placebo and caffeine
t = -2.5273, df = 21.999, p-value = 0.01918
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.4894368 -0.3438965
sample estimates:
mean of x mean of y
 24.25000  26.16667
```

Problem 1: the data, descriptives and the inferential test

```
prob1 #this shows us the data
   placebo caffeine
1       24       24
2       25       29
3       27       26
4       26       23
5       26       25
6       22       28
7       21       27
8       22       24
9       23       27
10      25       28
11      25       27
12      25       26
> dim(prob1) #how many rows and columns?
[1] 12  2
> describe(prob1)  #descriptive statistics
         vars  n  mean   sd median trimmed  mad min max range  skew kurtosis   se
placebo     1 12 24.25 1.86   25.0    24.3 1.48  21  27     6 -0.33    -1.33 0.54
caffeine    2 12 26.17 1.85   26.5    26.2 2.22  23  29     6 -0.22    -1.33 0.53
   #we specify that we have equal variances
> with(prob1, t.test(placebo,caffeine, equal.var=TRUE))

        Welch Two Sample t-test
data:  placebo and caffeine
t = -2.5273, df = 21.999, p-value = 0.01918
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.4894368 -0.3438965
sample estimates:
mean of x mean of y
 24.25000  26.16667
```

Problem 2: Is there a linear relationship?

First put the data into R, show them then describe them.

R code

```
#create a new data.frame
ie <- data.frame(Introversion=c(21,14,13,13,20,21,11,15,23,12,17,26),
        Spelling =c(31,33,39,24,35,37,36,20,46,31,44,44))
ie  # for a small data set, we can actually show the complete data
describe(ie) #basic descriptive statistics
```

```
   Introversion Spelling
1            21       31
2            14       33
3            13       39
4            13       24
5            20       35
6            21       37
7            11       36
8            15       20
9            23       46
10           12       31
11           17       44
12           26       44
```

```
             vars  n  mean  sd median trimmed  mad min max range  skew kurtosis   se
Introversion    1 12 17.17 4.9   16.0    16.9 5.93  11  26    15  0.33    -1.46 1.41
Spelling        2 12 35.00 7.9   35.5    35.4 6.67  20  46    26 -0.34    -0.99 2.28
```

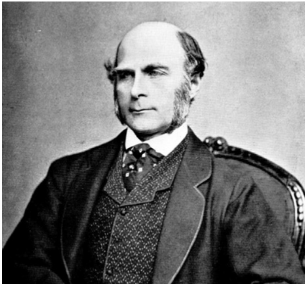
The correlation coefficient

1. Initial development by Francis Galton expressed as "reversion to mediocrity" (regression to the mean)
2. expressed in terms of interquartile range of x and y
3. Further developed by Karl Pearson as the products of two standard scores
4. Discussed by Charles Spearman in terms of similarities of ranks
5. Covariance of x and y (deviation scores)
   $\sigma_{xy} = \frac{\Sigma_{xy}}{N-1} = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{N-1}$

Francis Galton 1822-1911

Francis Galton (1822-1911) was among the most influential psychologists of the 19th century. He did pioneering work on the correlation coefficient, behavior genetics and the measurement of individual differences. He introspectively examined the question of free will and introduced the lexical hypothesis to the study of personality and character. In addition to psychology, he did pioneering work in meteorology and introduced the scientific use of fingerprints. Whenever he could, he counted.

http://personality-project.org/revelle/publications/galton.pdf



Karl Pearson 1857-1936

Carl (Karl) Pearson was among the most influential statisticians of the early 20th century. Founder of the statistics department at University College London. He developed the Pearson Product Moment Correlation Coefficient, its special case the ϕ coefficient, and the tetrachoric correlation. Major behavior geneticist and eugenicist.

Charles Spearman 1863-1945

Charles Spearman (1863-1945) was the leading psychometrician of the early 20th century. His work on the classical test theory, factor analysis, and the g theory of intelligence continues to influence psychometrics, statistics, and the study of intelligence. More than 100 years after their publication, his most influential papers remain two of the most frequently cited articles in psychometrics and intelligence. http://personality-project.org/revelle/publications/spearman.pdf

Bivariate Regression

X $\qquad$ Y $\qquad$ ϵ

$$y = \hat{y} + \epsilon = \beta_{y.x} x + \epsilon$$

$$\beta_{y.x} = \frac{\sigma_{xy}}{\sigma_x^2}$$

$$\epsilon = y - \hat{y}$$

$$\sum(\epsilon^2) = \sum(y - \hat{y})^2 = \sum(y - \beta_{y.x} x)^2 = \sum(y^2 - 2y\beta_{y.x} x + (\beta_{y.x} x)^2)$$

Minimize $\sum(\epsilon^2) w.r.t.\beta => \frac{d(\epsilon^2)}{d\beta} = 0 => -2\sigma_{xy} + 2\beta_{y.x}\sigma_x^2 = 0 =>$

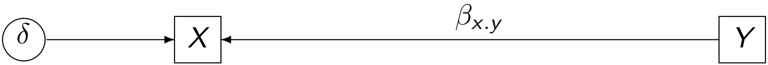
$$\beta_{y.x} = \frac{\sigma_{xy}}{\sigma_x^2}$$

Bivariate Regression

δ X Y ϵ

$$y = \hat{y} + \epsilon = \beta_{y.x}x + \epsilon$$

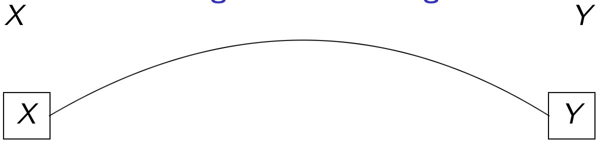
$$\beta_{y.x} = \frac{\sigma_{xy}}{\sigma_x^2}$$



$$x = \hat{x} + \delta = \beta_{x.y}y + \delta$$

$$\beta_{y.x} = \frac{\sigma_{xy}}{\sigma_y^2}$$

Bivariate Correlation is the geometric average of the two regressions



$x = \hat{x} + \delta = \beta_{x.y} y + \delta$

$y = \hat{y} + \epsilon = \beta_{y.x} x + \epsilon$

$\beta_{y.x} = \frac{\sigma_{xy}}{\sigma_y^2}$

$\beta_{y.x} = \frac{\sigma_{xy}}{\sigma_x^2}$

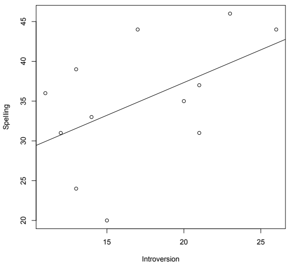
$$r_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2 \sigma_y^2}}$$

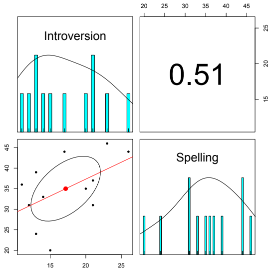
$$r_{xy} = \sigma_{z_x z_y} \text{(the covariance of standard scores)}$$

Linear relationships

This is asking if Spelling varies by Introversion and vice versa. This is the Pearson Product Moment *correlation coefficient*. There are multiple ways of seeing the relationship. A useful technique is to view the scatter plat as well as the correlation and the basic descriptive statistics. We can plot the data and fit a best fitting line, or we can use `pairs.panels` and ask for the linear regression.

R code
```
plot(Spelling ~ Introversion, data=ie)
abline(lm(Spelling ~ Introversion, data=ie))
pairs.panels(ie,lm=TRUE)
```





"Hand" calculate these values

R code

```
deviation.scores <- scale(ie, scale = FALSE) # find devation scores
cross.prod <- deviation.scores[,1] * deviation.scores[,2]
covariance <- sum(cross.prod/11)
ie.var <-  colSums(ie.scaled^2)/11
r <- covariance/sqrt(ie.var[1] * ie.var[2])
#show the results
ie.var ; covariance; r
# or use built in functions to find variance/covariance and r
var(ie)
 cor(ie)
```

```
ie.var ; covariance; r
Introversion     Spelling
    23.96970     62.36364
[1] 19.72727
Introversion
   0.5102348

# or use built in functions to find variance/covariance and r
> var(ie)
             Introversion Spelling
Introversion     23.96970 19.72727
Spelling         19.72727 62.36364
>  cor(ie)
             Introversion  Spelling
Introversion    1.0000000 0.5102348
Spelling        0.5102348 1.0000000
```

Inferential tests

1. How likely are correlations or regressions this big or bigger to happen by chance?
2. The `cor.test` will test a single correlation.
3. The `corr.test` will test one or more correlations.
4. The `lm` function returns the regression and gives probabilities, as will `setCor`

Four different ways of testing the correlation

R code

```
with(ie, cor.test(Spelling, Introversion))
corr.test(ie)  #test all of them
summary(lm(Spelling ~ Introversion, data=ie))
setCor(Spelling ~ Introversion, data=ie, std=FALSE)
```

```
> with(ie, cor.test(Spelling, Introversion))

        Pearson's product-moment correlation

data:  Spelling and Introversion
t = 1.8761, df = 10, p-value = 0.0901
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.09002976  0.83857967
sample estimates:
      cor
0.5102348

> corr.test(ie)  #test all of them
Call:corr.test(x = ie)
Correlation matrix
             Introversion Spelling
Introversion         1.00     0.51
Spelling             0.51     1.00
Sample Size
[1] 12
Probability values (Entries above the diagonal are adjusted for multiple tests.)
             Introversion Spelling
Introversion         0.00     0.09
Spelling             0.09     0.00
 To see confidence intervals of the correlations, print with the short=FALSE option
```

lm and setCor output

R code

```
summary(lm(Spelling ~ Introversion, data=ie))
setCor(Spelling ~ Introversion, data=ie, std=FALSE)
```

```
summary(lm(Spelling ~ Introversion, data=ie))
Call:
lm(formula = Spelling ~ Introversion, data = ie)
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   20.8717     7.8064   2.674   0.0233 *
Introversion   0.8230     0.4387   1.876   0.0901 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.123 on 10 degrees of freedom
Multiple R-squared:  0.2603,        Adjusted R-squared:  0.1864
F-statistic:  3.52 on 1 and 10 DF,  p-value: 0.0901

> setCor(Spelling ~ Introversion, data=ie, std=FALSE)
Call: setCor(y = Spelling ~ Introversion, data = ie, std = FALSE)

 DV =  Spelling
             slope   se    t     p lower.ci upper.ci   VIF
(Intercept)  20.87 7.81 2.67 0.023     3.48    38.27 14.41
Introversion  0.82 0.44 1.88 0.090    -0.15     1.80 14.41

Residual Standard Error =  7.12  with  10  degrees of freedom

 Multiple Regression
           R   R2  Ruw R2uw Shrunken R2 SE of R2 overall F df1 df2      p
Spelling 0.51 0.26 0.36 0.13        0.19     0.16      3.52   1  10 0.0901
>
>
```

Still another investigator believes that spelling performance is a function of the interaction of caffeine and time of day. She administers 0 or 200 mg of caffeine to subjects at 9 am and 9 pm. If the results are as below (Table 2) what statistical test should be applied to test her hypothesis?

Table: Time of day, caffeine, and spelling performance

9am	9 am	9pm	9pm
0 mg	200 mg	0 mg	200 mg
26	27	28	24
27	30	27	23
25	28	25	25
22	32	25	21
27	25	31	23
23	29	32	21
21	31	25	25
28	28	32	21
21	28	26	26
23	26	25	22
20	29	27	23
23	31	26	26

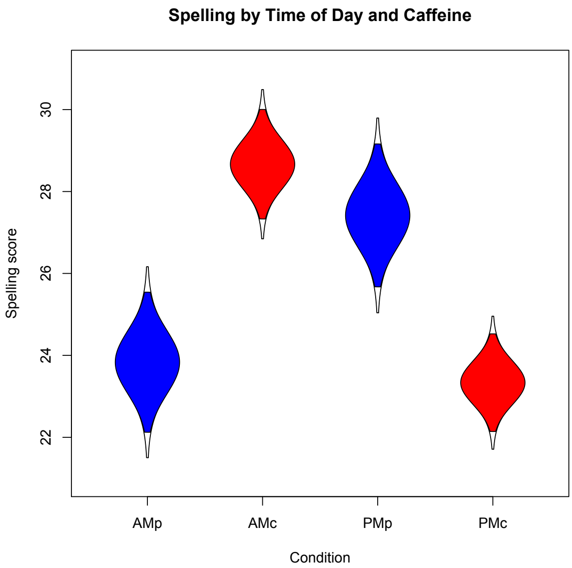
Multiple Independent Variables and the analysis of variance

R code

```
#show the data
prob3
d <- describe(prob3)  #we want to save the statistics for a graphic
d
 error.bars(stats=d,col=rep(c('blue',"red"),2),ylab="Spelling score",xlab="Cond
      main="Spelling by Time of Day and Caffeine")
```

```
 AMp AMc PMp PMc
1   26  27  28  24
2   27  30  27  23
3   25  28  25  25
4   22  32  25  21
5   27  25  31  23
6   23  29  32  21
7   21  31  25  25
8   28  28  32  21
9   21  28  26  26
10  23  26  25  22
11  20  29  27  23
12  23  31  26  26
> describe(prob3)
    vars  n  mean   sd median trimmed  mad min max range  skew kurtosis   se
AMp    1 12 23.83 2.69   23.0    23.8 2.97  20  28     8  0.16    -1.60 0.78
AMc    2 12 28.67 2.10   28.5    28.7 2.22  25  32     7 -0.08    -1.19 0.61
PMp    3 12 27.42 2.75   26.5    27.2 2.22  25  32     7  0.71    -1.25 0.79
PMc    4 12 23.33 1.87   23.0    23.3 2.97  21  26     5  0.09    -1.59 0.54
```

Prob 3: Graphical display



Analysis of variance as a generalization of the t-test

1. If we have multiple Independent Variables, we can ask what is the effect of each one separately, and both combined.
2. Originally developed by R.A. Fisher to test the effectiveness of various fertilizers and seeds on yield.
3. Applied in experimental psychology to test whether the effects of one variable depend upon (interact with) the effects of another variable.
4. For two IVs (A and B) this produces 3 tests: The effect of A, the effect of B and the interactive effect of A and B.

Using the aov function requires a bit of recoding of our data

R code

```
#we form a new data frame that is in "long" format (the data are strung out)
ps.df<- data.frame(stack(prob3),Time=c(rep(0,24),rep(1,24)),
    Drug=c(rep(0,12),rep(1,12),rep(0,12),rep(1,12)))
ps.df[c(1:3,10:14,22:26,34:38,46:48),]     #show the data
```

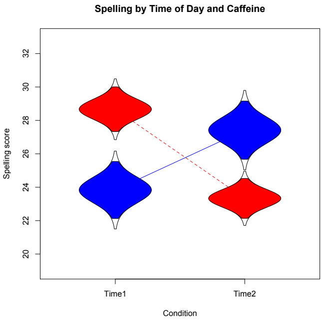
```
ps.df<- data.frame(stack(prob3),Time=c(rep(0,24),rep(1,24)),
   Drug=c(rep(0,12),rep(1,12),rep(0,12),rep(1,12)))
ps.df[c(1:3,10:14,22:26,34:38,46:48),]  #show the data
   values ind Time Drug
1      26 AMp    0    0
2      27 AMp    0    0
3      25 AMp    0    0
10     23 AMp    0    0
11     20 AMp    0    0
12     23 AMp    0    0
13     27 AMc    0    1
14     30 AMc    0    1
22     26 AMc    0    1
23     29 AMc    0    1
24     31 AMc    0    1
25     28 PMp    1    0
26     27 PMp    1    0
34     25 PMp    1    0
35     27 PMp    1    0
36     26 PMp    1    0
37     24 PMc    1    1
38     23 PMc    1    1
46     22 PMc    1    1
47     23 PMc    1    1
48     26 PMc    1    1
```

The aov function

R code

```
summary(aov(values ~ Time * Drug,data=ps.df))
error.bars.by(values ~ Time * Drug,data=ps.df,col=rep(c('blue',"red"),2),
     ylab="Spelling score",xlab="Condition",
       main="Spelling by Time of Day and Caffeine")
```

```
            Df Sum Sq Mean Sq F value   Pr(>F)    
Time         1   9.19    9.19   1.618    0.210    
Drug         1   1.69    1.69   0.297    0.588    
Time:Drug    1 238.52  238.52  41.994 6.63e-08 ***
Residuals   44 249.92    5.68                     
```



1. This is a nice example of a cross over interaction
2. There is no main efffect for Time, nor for Drug
3. But Drug helps spelling in the morning and hurts it in the evening

Another experimenter wants to test the hypothesis that gender is related to interest in football. 100 subjects (50 male and 50 female) are asked whether or not they watched a recent football game. The results are in Table 3

Table: Gender differences in football interest

	Watched	Did not watch
Male	30	20
Female	20	30

What statistical test should be applied to determine if there is a relationship between gender and watching the football game?

The χ^2 test of association

1. There are a number of ways to address this problem. One is to compare the observed frequencies with the expected frequencies and test for differences using the χ^2 test.
2. A simple alternative is to find the ϕ correlation coefficient between watching football and gender.
3. χ^2 is frequently used for nominal data and their resulting frequencies.

χ^2

1. Find the expected cell frequencies given the row and column marginal frequencies
2. Compare these expected values to the observed values.
3. In our particular case, the rows and columns suggest 50% for each row and column and thus 25% for each cell.
4. Expected in every cell is thus 25
5. $\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(20-25)^2}{25} + \frac{(30-25)^2}{25} + \frac{(30-25)^2}{25} + \frac{(20-25)^2}{25} = 4$

```
chisq.test(prob4,correct=FALSE)
        Pearson's Chi-squared test
data:  prob4
X-squared = 4, df = 1, p-value = 0.0455

but
chisq.test(prob4)
Pearson's Chi-squared test with Yates' continuity correction

data:  prob4
X-squared = 3.24, df = 1, p-value = 0.07186
```

Problem 5: t-tests with independent groups

A professor believes that taking statistics increases one's ability to reason analytically. To test this hypothesis, she develops a test of reasoning and gives it to two sets of students. Those who have just started a statistics course and those who have just finished a statistics course. The results are shown in Table 4

Table: The effect of taking a statistics course on reasoning analytically.

before	after
12	15
11	23
15	17
14	22
11	18
10	17
11	21
12	21
18	16
17	17
13	23
16	18

What test should be applied to these data to test her hypothesis?

Problem 5 is just two independent groups

1. 2 independent groups. Do they differ? Do a t-test

R code

```
teaching <- data.frame(pre=c(12,11,15,14,11,10,11,12,18,17,13,16),
                     post = c(15,23,17,22,18,17,21,21,16,17,23,18))
 describe(teaching)
 with(teaching, t.test(pre,post))
```

```
 describe(teaching)
    vars  n  mean   sd median trimmed  mad min max range skew kurtosis   se
pre     1 12 13.33 2.64   12.5    13.2 2.22  10  18     8 0.43    -1.40 0.76
post    2 12 19.00 2.83   18.0    19.0 3.71  15  23     8 0.20    -1.68 0.82
> with(teaching, t.test(pre,post))

	Welch Two Sample t-test

data:  pre and post
t = -5.0735, df = 21.896, p-value = 4.47e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -7.983620 -3.349713
sample estimates:
mean of x mean of y
 13.33333  19.00000
```

Correlated t-tests

1. Another professor has the same hypothesis, but decides to use a pre-post design. That is, each student takes the reasoning test twice, once before and once after the class.
2. This is a t-with paired subjects.
3. By using a within subjects design, the study has much more power to detect differences.
4. However, the sample size is reduced from 24 to 12 which reduces the power.

R code

```
with(teaching, t.test(pre,post, paired=TRUE))
```

```
Paired t-test

data:  pre and post
t = -4.363, df = 11, p-value = 0.001131
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -8.525295 -2.808038
sample estimates:
mean of the differences
              -5.666667
```

Probability theory

1. If a test is normally distributed and has a mean of 100 and a standard deviation of 15, then what percentage of students would you expect to have scores of 100 or greater?
2. With the same assumptions, what percentage of students would you expect to have scores greater than 115?
   - This is just knowing the distribution of a normal curve.
   - 50% will have a score above the mean
   - 1- .84 = .16 will be above 1 standard deviation

```
pnorm(1)     #in terms of mean 0 and sd = 1
[1] 0.8413447
#or
pnorm(115,100,15) #specify the score, the mean, the sd
[1] 0.8413447
```

3. If you flip a fair coin 10 times, how often would you expect to observe at least 8 heads? This is asking for the distribution of a binomial with a particular probability.

```
coin.df <- data.frame(heads=0:10,frequency =dbinom(0:10,10,.5) )   #make up a probability table
tcoin.df <- t(coin.df[,2])   #transpose it
colnames(tcoin.df) <- coin.df[,1]  #provide names
round(tcoin.df,2)  #round and display
     0    1    2    3    4    5    6    7    8    9 10
[1,] 0 0.01 0.04 0.12 0.21 0.25 0.21 0.12 0.04 0.01  0
```