1. The following data are measurements of the smoothness of five types of paper obtained in four different laboratories. Carry out a Friedman test of the null hypothesis of no systematic difference amongst the laboratories.

<table>
<thead>
<tr>
<th>Type</th>
<th>Laboratory</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>38.7 39.2 34.0 34.0</td>
</tr>
<tr>
<td>II</td>
<td>41.5 39.3 35.0 34.8</td>
</tr>
<tr>
<td>III</td>
<td>43.8 39.7 39.0 34.8</td>
</tr>
<tr>
<td>IV</td>
<td>44.5 41.8 40.0 35.4</td>
</tr>
<tr>
<td>V</td>
<td>45.5 41.8 43.0 37.2</td>
</tr>
</tbody>
</table>

**Solution:** The hypothesis to be tested is

\[ H : \text{No systematic differences among the labs} \quad \text{vs} \quad H_A : \text{There are systematic differences among the labs}. \]

The Friedman test depends on the ranks \( r_{ij} \) of the \( i \)-th block(type), which is given in the brackets below:

<table>
<thead>
<tr>
<th>Type</th>
<th>Laboratory</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>38.7 (3) 39.2 (4) 34.0 (1.5) 34.0 (1.5)</td>
</tr>
<tr>
<td>II</td>
<td>41.5 (4) 39.3 (3) 35.0 (2) 34.8 (1)</td>
</tr>
<tr>
<td>III</td>
<td>43.8 (4) 39.7 (3) 39.0 (2) 34.8 (1)</td>
</tr>
<tr>
<td>IV</td>
<td>44.5 (4) 41.8 (3) 40.0 (2) 35.4 (1)</td>
</tr>
<tr>
<td>V</td>
<td>45.5 (4) 41.8 (2) 43.0 (3) 37.2 (1)</td>
</tr>
</tbody>
</table>

\[
\sum_{i=1}^{5} r_{ij} = 19 \quad 15.0 \quad 10.5 \quad 5.5
\]
\[
\sum_{i=1}^{5} r_{ij}^2 = 73 \quad 47 \quad 23.25 \quad 6.25
\]

This yields that (note \( \bar{r} = (4 + 1)/2 = 2.5 \))

\[
\bar{r}_{\cdot j} = 3.8, \ 3.0, \ 2.1, \ 1.1,
\]
\[
SS_t = 5 \sum_{j=1}^{4} (\bar{r}_{\cdot j})^2 - 5 \times 4 \times (\bar{r})^2 = 20.3
\]
\[
SS_e = \frac{1}{5(4-1)} \left( \sum_{i=1}^{5} \sum_{j=1}^{4} r_{ij}^2 - rs(\bar{r})^2 \right) = 1.633
\]
\[
Q = SS_t/SS_e = 12.43.
\]
The corresponding $p$-value is given by

$$p \approx P(\chi^2 \geq 12.43) = 0.006049655.$$  

**Conclusion:** There is very strong evidence that there are systematic differences among the labs.

2. In the case of no ties, show that the Friedman test statistic is

$$Q = \frac{12r}{s(s+1)} \sum_{j=1}^{s} (\bar{r}_{j})^2 - 3r(s+1).$$

Hint: Recall the notation in Lect. 22 and also note that $\bar{r} = (1+s)/2$ and in the case of no ties, for each $i$,

$$\sum_{j=1}^{s} r_{ij}^2 = \sum_{i=1}^{s} i^2 = \frac{1}{6} s(s+1)(2s+1).$$

**Solution:** In the case of no ties, for each fixed $i$, $x_{ij}$ has a different rank for $j = 1, 2, ..., s$. This yields that, for each fixed $i$,

$$\sum_{j=1}^{s} r_{ij} = \sum_{j=1}^{s} j = \frac{1}{2} s(s+1),$$

$$\sum_{j=1}^{s} r_{ij}^2 = \sum_{j=1}^{s} j^2 = \frac{1}{6} s(s+1)(2s+1).$$

Hence, $\bar{r} = (s + 1)/2$,  

$$SS_t = r \sum_{j=1}^{s} (\bar{r}_{j} - \bar{r})^2 = r \sum_{j=1}^{s} (\bar{r}_{j})^2 - rs(\bar{r})^2,$$

$$= r \sum_{j=1}^{s} (\bar{r}_{j})^2 - \frac{1}{4} rs(s+1)^2.$$

$$SS_e = \frac{1}{r(s-1)} \sum_{i=1}^{r} \sum_{j=1}^{s} (r_{ij} - \bar{r})^2$$

$$= \frac{1}{r(s-1)} \left( \sum_{i=1}^{r} \sum_{j=1}^{s} r_{ij}^2 - rs(\bar{r})^2 \right)$$

$$= \frac{1}{6(s-1)} s(s+1)(2s+1) - \frac{s}{4(s-1)} (s+1)^2$$

$$= \frac{1}{12} s(s+1).$$

It follows that

$$Q = \frac{SS_t}{SS_e} = \frac{12r}{s(s+1)} \sum_{j=1}^{s} (\bar{r}_{j})^2 - 3r(s+1).$$
1. How does the frequency that a supermarket product is promoted at a discount affect
the price that customers expect to pay for the product? Does the percent reduction
also affect expectation? Results of a study in Australia are given in a data frame
called price.df. The number of promotions were 1, 3, 5, 7 and the percent of reduction
were 10, 20, 30 and 40. The response is the expected price (obtained from a customer
survey) of a supermarket product. Perform a detailed two-way ANOVA for this data
by following Lect.24.

(a) Set a 2 by 2 graphic window and prepare your data for analysis by typing
attach(price.df).

(b) Obtain group means of the data, say price.mean.

(c) Plot the group means and comment on the results.

(d) Compute the overall mean and the marginal means of the group means. Store
the results as overall.mean, price.row and price.col respectively.

(e) Compute group sample variances and store the results as price.var.

(f) Find the number m of replicates.

(g) Compute the TMS, BMS, IMS and Residual MS ($s^2$) on the two way ANOVA
table and store the results as TMS, BMS, IMS and RMS respectively.

(h) Compute the p-values corresponding to the different hypotheses for this data, and
comment on the results.

(i) Check your results by using the built-in aov function. Obtain a two-way table
as well as a normal-quantile plot of the residuals. Comment on the assumptions
required for the two-way ANOVA.

Solution:

```r
# (a)
par(mfrow=c(2,2))
attach(price.df)

# (b)
price.mean<- tapply(pri,list(discount, promotion),mean))

> price.mean
      1      3      5      7
10% 4.423 4.284 4.058 3.780
20% 4.225 4.097 3.890 3.760
30% 4.689 4.524 4.251 4.094
```
40% 4.920 4.756 4.393 4.269

# (c)

interaction.plot(discount,promotion, pri)
interaction.plot(promotion,discount, pri)

# From the plot of group means we see that
# A: expected price is decreasing as the promotion number is increasing
# B: expected price is increasing with % of discount.
# C: It seems no interactions between the promotion number
# and the discount %

# (d)
overall.mean<-mean(pri)
> overall.mean
[1] 4.275813

price.row<-apply(price.mean,1,mean)
> price.row
  10% 20% 30% 40%
4.13625 3.993 4.3895 4.5845

price.col<-apply(price.mean,2,mean)
> price.col
   1  3  5  7
4.56425 4.41525 4.148 3.97575

# (e)
price.var<- tapply(pri,list(discount,promotion),var)

> price.var
         1         3         5         7
10% 0.03413444 0.04162667 0.03097333 0.04595556
20% 0.14869444 0.05504556 0.02653333 0.06853333
30% 0.05432111 0.07329333 0.07014333 0.05796000
40% 0.02311111 0.05900444 0.07211222 0.07285444

# (f)
s<-4, r<-4, m<-length(pri)/(r*s)

# (g)
### (h)
*The p-value for discount*

```r
tms <- TSS / (s - 1)
> tms
[1] 2.786834

bms <- BSS / (r - 1)
> bms
[1] 2.768979

RMS <- RSS / (r * s * (m - 1))
> RMS
[1] 0.05839354
```

### (i)
*The p-value for promotion*

```r
p1 <- 1 - pf(TMS/RMS, r - 1, r*s*(m - 1))
> p1
[1] 0

p2 <- 1 - pf(BMS/RMS, s - 1, r*s*(m - 1))
> p2
[1] 0

p3 <- 1 - pf(IMS/RMS, (r - 1)*(s - 1), r*s*(m - 1))
> p3
[1] 0.9121002
```

### (i)
*Checking results*
price.aov<-aov(pri~ promotion*discount)
> summary(price.aov)

|                | Df | Sum of Sq | Mean Sq | F Value | Pr(>|F|) |
|----------------|----|-----------|---------|---------|---------|
| promotion      | 3  |  8.360502 |  2.786834 | 47.72504 | 0.0000000 |
| discount       | 3  |  8.306937 |  2.768979 | 47.41927 | 0.0000000 |
| promotion:discount | 9 |  0.230586 |  0.025621 |  0.43876 | 0.9121002 |
| Residuals      | 144|  8.408670 |  0.058394 |         |         |

> boxplot(price.aov$resid)
> qqnorm(price.aov$resid)
> qqline(price.aov$resid)

# The boxplot looks symmetric and the qq-plot looks quite linear
# Hence it is reasonable to believe the data satisfies the conditions
# required for two-way ANOVA.