Supplementary exercise 13.31 of IPS7e

Data: A completely randomized 4×4 factorial trial on how promotions affect consumer expectations. Students were subjected to promotions during a 10-week period, at the end of which their expected price for the product in question was recorded. The factors affecting the type of promotions received were the number of promotions (values $1, 3, 5, 7$) and the percent of discount (values $10\%, 20\%, 30\%, 40\%$). There were 10 replicates for each of the 16 treatments (= combinations of the two factors).

Model: The statistical model of interest is two-way ANOVA model:

$$\begin{aligned} X_{ijk} &= \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \quad \text{or} \\ \text{Price} &= \text{Promotion} + \text{Discount} + \text{Promotion} * \text{Discount} + \text{Error}, \end{aligned}$$

where

- μ = overall mean,
- α_i = effect of promotion group i,
- β_j = effect of discount group j,
- γ_{ij} = interaction effect of (promotion,discount) group (i, j),
- ε_{ijk} = error for obs. (i, j, k), assumed to be i.i.d. and $\mathrm{N}(0, \sigma)$.

(a) The requested descriptive statistics and the graph can be obtained as part of the two-way ANOVA analysis, but it is also fine to look at them as a preliminary step for the analysis. We usually get descriptive statistics from the `Basic Statistics` menu, but in order to display them in a tabular format (which is easier to look at), we need to use the `Tables` menu. The graphical display of interest is the interaction plot, available directly in the `ANOVA` menu.

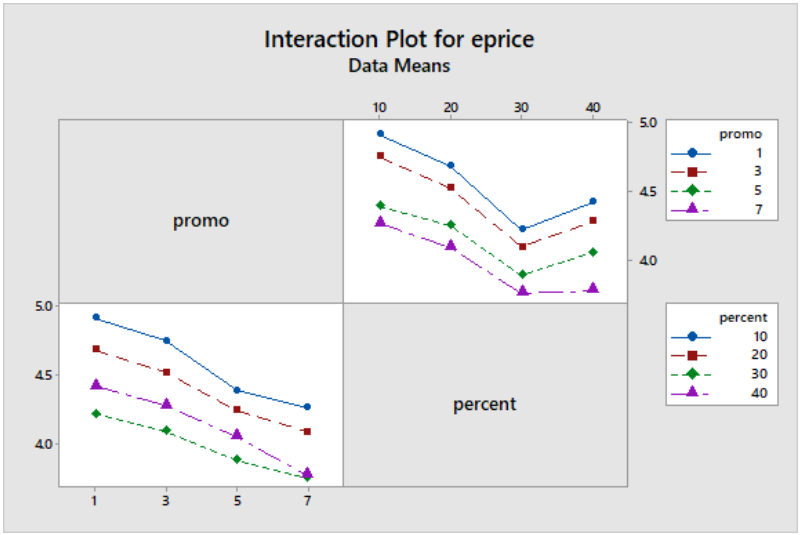
```
MTB > Table 'promo' 'percent';
SUBC>   Layout 1 1;
SUBC>   DMissing 'promo' 'percent';
SUBC>   Means 'eprice';
SUBC>   StDev 'eprice';
SUBC>   Counts.
MTB > Interact 'promo' 'percent';
SUBC>   Response 'eprice';
SUBC>   Full.
```

Tabulated Statistics: promo, percent

Rows: promo Columns: percent

	10	20	30	40	All
1	4.920	4.689	4.225	4.423	4.564
	0.1520	0.2331	0.3856	0.1848	0.3621
	10	10	10	10	40
3	4.756	4.524	4.097	4.284	4.415
	0.2429	0.2707	0.2346	0.2040	0.3406
	10	10	10	10	40
5	4.393	4.251	3.890	4.058	4.148
	0.2685	0.2648	0.1629	0.1760	0.2887
	10	10	10	10	40
7	4.269	4.094	3.760	3.780	3.976
	0.2699	0.2407	0.2618	0.2144	0.3225
	10	10	10	10	40
All	4.584	4.390	3.993	4.136	4.276
	0.3518	0.3376	0.3196	0.3099	0.3990
	40	40	40	40	160

Cell Contents
eprice : Mean
eprice : Standard deviation
Count



Comments:
The means for the promotion groups show that the expected price decreases with the number of promotions, almost linearly. Also, the expected price decreases with the discount, except for 30% and 40% where the values are interchanged. The lines in the plot are almost parallel, indicating that the interaction is small relative to the main effects.

(b) We start with the simpler `Balanced ANOVA` menu, including also residual plots although this menu does not offer standardized residuals (so we have to use the raw residuals). However, because the design is balanced, all residuals will have the same standard error, so the standardized residuals are simply a uniform scaling of the raw residuals, and the patterns will be exactly the same.

```
MTB > ANOVA 'eprice' = promo percent promo* percent;
SUBC>   GFourpack.
```

ANOVA: eprice versus promo, percent

Factor Information

Factor	Type	Levels	Values
promo	Fixed	4	1, 3, 5, 7
percent	Fixed	4	10, 20, 30, 40

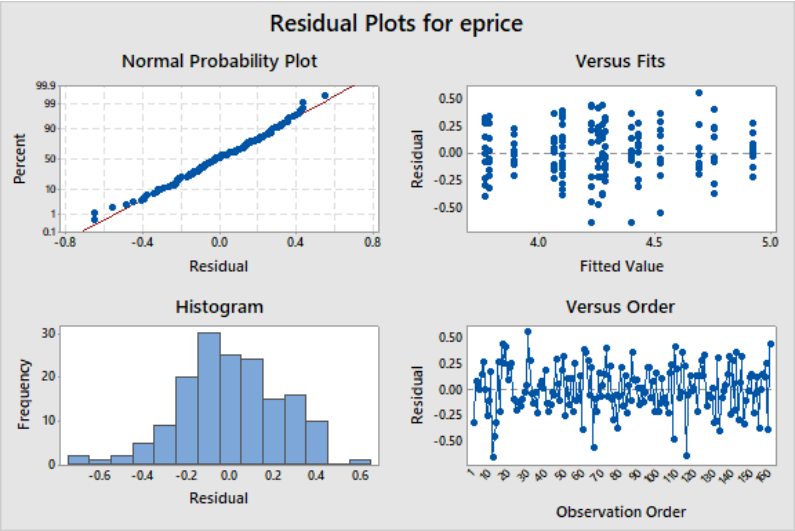
Analysis of Variance for eprice

Source	DF	SS	MS	F	P
promo	3	8.3605	2.78683	47.73	0.000
percent	3	8.3069	2.76898	47.42	0.000
promo*percent	9	0.2306	0.02562	0.44	0.912
Error	144	8.4087	0.05839		
Total	159	25.3067			

Model Summary

S	R-sq	R-sq(adj)
0.241648	66.77%	63.31%

Residual Plots for eprice



Comments:
The ANOVA table shows indeed that the interaction is small and clearly non-significant, whereas the two main effects are strongly significant. The residual plots show no reasons for concern, and a normality test for the residuals is clearly non-significant (not shown; one would need to first store the residuals). The standard deviations are similar across the 16 groups and only mildly violate our guideline (for 1-Way ANOVA): $s_{\max}/s_{\min} = 0.3856/0.1520 = 2.5$. With so many groups, the rule becomes quite strict, and one may suspect that it is unnecessarily strict in this case. One way to formally assess that is by the tests for equal variances (also available in the `ANOVA` menu); indeed both tests give clearly non-significant P-values (0.347 and 0.416, not shown). We conclude that the equal variances assumption poses no reason for concern with these data.

Because of the non-significant interaction, our continued analysis will focus on each of the main effects separately, as estimated in the two-way ANOVA model. We could remove the interaction from the model, but as DFE is already very large there is no pressing need to do so. Without any prespecified hypotheses, the natural way to continue analysis is by multiple comparisons, for which we either have to do a manual LSD-calculation, or reanalyze using the `General Linear Model` menu.

We start with the LSD-calculation. Both factors (`promo` and `discount`) have 4 levels, so for both we need to take into account a total of $4 \cdot 3/2 = 6$ multiple comparisons. For the Bonferroni method, the adjusted significance level for individual comparisons becomes $0.05/6$, and our (adjusted) t^* therefore is the percentile from a $t(\text{DFE})$ distribution corresponding to a tail probability of $0.025/6 = 0.0417$; Minitab gives this value as 2.67491.

```
MTB > InvCDF .00417;
SUBC>   T 144.
```

Inverse Cumulative Distribution Function

Student's t distribution with 144 DF

P(X ≤ x)	x
0.00417	-2.67491

With this value, and using the fact that all means of interest (for either `promo` or `discount`) are across 40 observations, we get

$$\text{LSD}(.99583) = t^* \cdot s\sqrt{2/40} = 2.67491 \cdot 0.2416\sqrt{2/40} = 0.1445.$$

The LSD comparisons show that all levels of the main effects are significantly different at an overall (not invidual) 5% signigicance level, except the 30% and 40% discount groups where the difference is 0.143 and hence just short of the LSD threshold for significance.

Next we show Minitab output from the `Comparisons` menu (with the Bonferroni method) after first having fitted the model in the `General Linear Model` menu (note that you have to include the interaction in the `Model` submenu because it will not be included by default).

```
MTB > GLM;
SUBC>   Response 'eprice';
SUBC>   Nodefault;
SUBC>   Categorical 'promo' 'percent';
SUBC>   Terms promo percent
        promo*percent;
SUBC>   TMethod;
SUBC>   TAnova;
SUBC>   TSummary;
SUBC>   TCoefficients;
SUBC>   TEquation;
SUBC>   TFactor;
SUBC>   TDiagnostics 0.
MTB > Compare 'eprice';
SUBC>   Pairwise promo percent;
SUBC>     Bonferroni;
SUBC>   NoDefault;
SUBC>   TGrouping;
SUBC>   TMTest.
```

Comparisons for eprice

Bonferroni Pairwise Comparisons: promo

Grouping Information Using the Bonferroni Method and 95% Confidence

promo	N	Mean	Grouping
1	40	4.56425	A
3	40	4.41525	B
5	40	4.14800	C
7	40	3.97575	D

Means that do not share a letter are significantly different.

Bonferroni Simultaneous Tests for Differences of Means

Difference of promo Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
3 - 1	-0.1490	0.0540	(-0.2936, -0.0044)	-2.76	0.039
5 - 1	-0.4163	0.0540	(-0.5608, -0.2717)	-7.70	0.000
7 - 1	-0.5885	0.0540	(-0.7331, -0.4439)	-10.89	0.000
5 - 3	-0.2673	0.0540	(-0.4118, -0.1227)	-4.95	0.000
7 - 3	-0.4395	0.0540	(-0.5841, -0.2949)	-8.13	0.000
7 - 5	-0.1722	0.0540	(-0.3168, -0.0277)	-3.19	0.011

Individual confidence level = 99.17%

Bonferroni Pairwise Comparisons: percent

Grouping Information Using the Bonferroni Method and 95% Confidence

percent	N	Mean	Grouping
10	40	4.58450	A
20	40	4.38950	B
40	40	4.13625	C
30	40	3.99300	C

Means that do not share a letter are significantly different.

Bonferroni Simultaneous Tests for Differences of Means

Difference of percent Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
20 - 10	-0.1950	0.0540	(-0.3396, -0.0504)	-3.61	0.003
30 - 10	-0.5915	0.0540	(-0.7361, -0.4469)	-10.95	0.000
40 - 10	-0.4482	0.0540	(-0.5928, -0.3037)	-8.30	0.000
30 - 20	-0.3965	0.0540	(-0.5411, -0.2519)	-7.34	0.000
40 - 20	-0.2532	0.0540	(-0.3978, -0.1087)	-4.69	0.000
40 - 30	0.1432	0.0540	(-0.0013, 0.2878)	2.65	0.054

Individual confidence level = 99.17%

Comments:
The Bonferroni comparisons show that all levels of the main effects are significantly different at an overall 5% level, except the 30% and 40% discount groups with a P-value of 0.054. We also note that there were 5 standardized residuals outside of $(-2, 2)$ (not shown), but this is actually slightly less than expected for a data set of this size.

(c) The analysis showed no interaction effect of Discount and Promotion on the expected price. Both Discount and Promotion had clearly significant effects. The expected price decreases with the number of promotions and is significantly different between all promotion groups. The expected price decreases when the discount increases from 10% to 20% and 30%, also significantly so. However, the expected price takes a jump upwards from 30% to 40% even if this jump is only close to significant. The value at 40% is still significantly less than those at 10% and 20%.