SPSS Analysis of Inferential Statistics

THE COMPUTER AND BIVARIATE INFERENTIAL STATISTICS

In the textbook, we emphasized the logic and uses of various statistical tests rather than computational formulas. Because computers are almost always used for statistical analysis, and because it is likely to be useful to know how to read a computer printout, we include examples of computer analyses for two statistical tests in this Supplement, using output from the Statistical Package for the Social Sciences (SPSS).[1]

We return to the example described in the Supplement to Chapter 17, which involved a randomized trial to test the effects of a special prenatal program for young low-income women. In Figure 1, we reproduce the screenshot of the SPSS data file for this example, which shows the data for seven variables (ID, GROUP, AGE, PRIORS, SMOKE, BWEIGHT, and REPEAT) for 15 mothers in the experimental group and 15 mothers in the control group. Given these data, let us test some hypotheses.

Hypothesis 1: *t*-Test

Our first research hypothesis is that experimental group infants have higher birthweights than control

group infants. The *t*-test for independent samples is used to test the hypothesis of mean group differences. The null and alternative hypotheses are:

$$H_0 : \mu \text{ experimental} = \mu \text{ control}$$
$$H_A : \mu \text{ experimental} \neq \mu \text{ control}$$

We used the following SPSS commands to run the *t*-test analysis for the variables GROUP (the independent variable) and BWEIGHT (the dependent variable):

**ANALYZE → COMPARE MEANS
→ INDEPENDENT SAMPLES T - TEST**

Figure 2 presents the SPSS printout for the *t*-test. Panel A presents some descriptive statistics (mean, standard deviation, and standard error of the mean) for the birthweight variable, separately for the two groups. The mean birthweight of the babies in the experimental (intervention) group is 107.5333 ounces, compared with 101.8667 ounces for those in the control group. The data are consistent with the research hypothesis—i.e., the average weight of babies in the intervention group is higher than that of controls. But is the difference attributable to the intervention, or could it reflect random fluctuations in this sample of young women?

Panel B of Figure 2 first presents results of **Levene's test** for equality of variances. An assumption underlying use of the *t*-test is that the

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons

	ID	GROUP	AGE	PRIORS	SMOKE	BWEIGHT	REPEAT
1	1	1	17	1	1	107	1
2	2	1	14	0	0	101	0
3	3	1	21	3	0	119	0
4	4	1	20	2	0	128	1
	5	1	15	1	1	89	0
6	6	1	19	0	1	99	0
7	7	1	19	1	0	111	0
8	8	1	18	1	1	117	1
9	9	1	17	0	0	102	1
	10	1	20	0	0	120	0
11	11	1	13	0	1	76	0
12	12	1	18	0	1	116	0
13	13	1	16	0	0	100	1
14	14	1	18	0	0	115	0
	15	1	21	2	1	113	0
16	16	2	19	0	0	111	1
17	17	2	21	1	0	108	0
18	18	2	19	2	1	95	0
19	19	2	17	0	1	99	0
	20	2	19	0	0	103	1
21	21	2	15	0	1	94	0
22	22	2	17	1	0	101	1
23	23	2	21	2	0	114	0
24	24	2	20	1	0	97	0
	25	2	18	0	1	99	1
26	26	2	18	0	1	113	0
27	27	2	19	1	0	89	0
28	28	2	20	0	0	98	0
29	29	2	17	0	0	102	0
30	30	2	19	1	1	105	0

NOTES:

GROUP:	Group status,	1 = Experimental group	2 = Control group
AGE:	Mother's age in years		
PRIORS:	Number of prior pregnancies		
SMOKE:	Mother's smoking status,	1 = Smokes	0 = Does not smoke
BWEIGHT:	Infant's birthweight, in ounces		
REPEAT:	Had repeat pregnancy within 18 months,	1 = Yes	0 = No

FIGURE 1 Fictitious dataset for intervention study with low-income pregnant adolescents (screenshot of an SPSS Data File).

A

Group statistics

	Treatment group	N	Mean	Std. Deviation	Std. Error Mean
Infant birthweight in ounces	Experimental	15	107.5333	13.37838	3.45428
	Control	15	101.8667	7.23944	1.86922

B

Independent samples test

		Levene Test for Equality of Variances		t-Test for Equality of Means						95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (two-tailed)	Mean Difference	Std. Error Difference		Lower	Upper
Infant birthweight in ounces	Equal variances assumed	4.370	.046	1.443	28	.160	5.66667	3.92760		−2.37865	13.71199
	Equal variances not assumed			1.443	21.552	.163	5.66667	3.92760		−2.48851	13.82185

FIGURE 2 SPSS *t*-test printout: Testing group differences in infants' birthweight.

population variances for the two groups are equal. In Panel A, we can see that the standard deviations (and thus the variances) are quite different, with substantially more variability among those in the experimental group ($SD = 13.38$) than among those in the control group ($SD = 7.24$). Levene's test tells us that the two variances are, in fact, significantly different (Sig. = .046).

Panel B then presents two rows of *t*-test information. The top row is for the **pooled variance *t*-test,** which is used when equality of variances can be assumed. Given the significantly different variances in this sample, however, we should use information in the second row, which uses a different (**separate variance *t*-test**) formula. The mean group difference in birthweights is 5.66667 ounces. The value of the *t* statistic is 1.443, and the two-tailed probability (Sig.) for the differences in group means is .163. This means that in about 16 samples out of 100, we could expect a mean difference in weights this large as a result of chance. Therefore, because a *p* of .163 is greater than $p < .05$, this is a nonsignificant result. We cannot conclude that the intervention was effective in improving the

birthweights of experimental group infants. Note that we cannot conclude that it was *not* effective, either. Failure to reject the null hypothesis does not provide evidence that the null is true.

The last two columns of Panel B show the 95% confidence intervals for the population mean difference. We can conclude with 95% confidence that the mean *difference* in birthweights for the population of young mothers exposed and not exposed to the intervention lies between −2.48851 ounces and +13.82185 ounces. Zero is included in this confidence interval, indicating the possibility that there are no group differences in the population. This is consistent with the fact that we were not able to reject the null hypothesis of equal means on the basis of the *t*-test.

We noted in the textbook that a power analysis can be used to estimate an effect size. In our example, the effect size estimate is as follows:

$$ES = (107.5333 - 101.8667) \div 10.955 = .52$$

The estimated effect size is the experimental mean minus the control mean, divided by the overall standard deviation (i.e., the *pooled SD* for

the two groups combined), which is 10.955 (the printout for this, from the Descriptives program of SPSS, is not shown). The obtained effect size of .52 is moderate. However, when we look at the power table in Table 18.6 of the textbook, we see that with a sample size of only 15 per group, our power to detect a true population difference is less (actually *far* less) than .60. This means that with a sample of only 15 per group, we had a very high risk of a Type II error. We can also see that with an effect size of .52, we would have needed about 60 mothers in each group to achieve a power of .80. This demonstrates the importance of doing a power analysis before undertaking a study.

Hypothesis 2: Pearson Correlation

Our second research hypothesis is as follows: Older mothers have babies of higher birthweight than younger mothers. In this case, both birthweight and maternal age are measured on the ratio scale, so the appropriate test statistic is Pearson's product-moment correlation. The hypotheses are:

$$H_0 : \rho \text{ birthweight} - \text{age} = .00$$
$$H_A : \rho \text{ birthweight} - \text{age} \neq .00$$

For this analysis, we used the following SPSS commands to examine the correlation between AGE and BWEIGHT:

ANALYZE → CORRELATE → BIVARIATE

The SPSS printout for the hypothesis test is presented in Figure 3. The correlation matrix shows, in row 1, the correlation of infant birthweight with infant birthweight and of birthweight with mother's age; and in row 2, the correlation of mother's age with infant birthweight and of age with age. In the shaded cell at the intersection of age and birthweight, we find three numbers. The first is the correlation coefficient ($r = .594$), which indicates a moderately strong positive relationship: the older the mother, the higher the baby's weight tended to be, consistent with the research hypothesis. The second number in the cell shows the probability that the correlation occurred by chance: Sig. (for significance level) = .001 for a two-tailed test. In other words, a relationship this strong would be found by chance alone in about 1 out of 1,000 samples of 30 young mothers. Therefore, the research hypothesis is accepted. The final number in the shaded cell is 30, the total sample size (N).

We can use these two variables (AGE and BWEIGHT) to illustrate a scatterplot—that is, a graphic representation of the relationship between these two continuous variables. To create a scatterplot, we used the following SPSS commands:

GRAPHS → LEGACY DIALOGS → SCATTER/DOT

The resulting scatterplot is shown in Figure 4. Consistent with the value of Pearson's r statistic, the pattern of "dots" in this diagram shows that there is a moderately high, positive relationship between maternal age and the infant's birthweight.

Correlations

		Infant birthweight in ounces	Mother's age
Infant birthweight in ounces	Pearson correlation	1.000	.594**[*]
	Sig. (two-tailed)		.001
	N	30.000	30
Mother's age	Pearson Correlation	.594**[**]	1.000
	Sig. (two-tailed)	.001	
	N	30	30.000

**Correlation is significant at the 0.01 level (two-tailed).

FIGURE 3 SPSS correlation matrix printout: Testing the relationship between maternal age and infant birthweight.
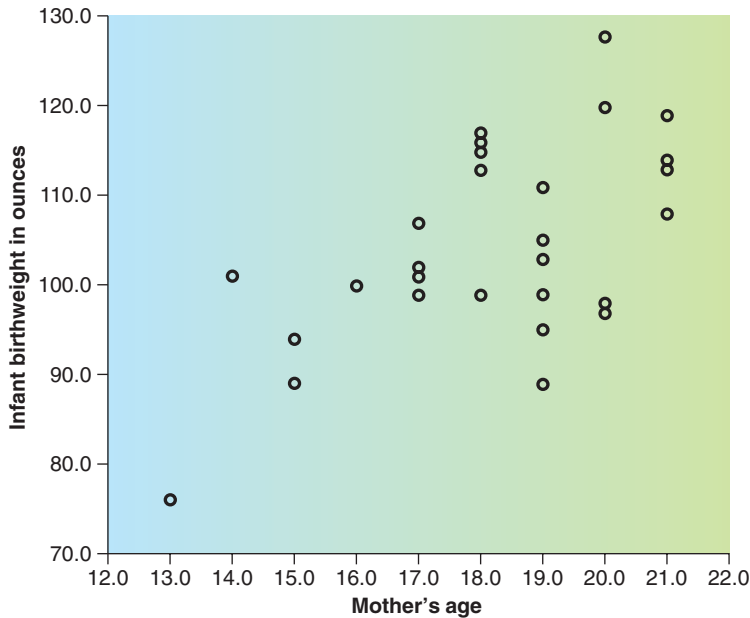
FIGURE 4 SPSS scatterplot printout: Maternal age and infant birthweight.

Hypothesis 3: Chi-Square Test

Our third hypothesis concerns the relationship between maternal age (AGE) and whether or not there was a repeat pregnancy within 18 months after the birth of the focal infant (REPEAT). We predict that younger mothers were more likely than older ones to have an early repeat pregnancy. One way to analyze this would be to run a *t*-test for independent groups, comparing the mean age of the mothers who did or did not have a repeat pregnancy.

We will use an alternative so that we can show output for a chi-square test. In this analysis, we categorized study participants as either being 17 years old or younger ($n = 10$) or 18 years or older ($n = 20$), by creating a new variable that we called AGEGRP.[2] Then we crosstabulated AGEGRP with REPEAT, using these commands:

ANALYZE → DESCRIPTIVE STATISTICS → CROSSTABS

The results are shown in Figure 5. Each of the four main cells in the 2 × 2 crosstabs table in Panel A has three pieces of information: the actual count (number of mothers in that cell), the *expected* count (the count expected if the null hypothesis were true), and the column percentage (number of observations in the cell, divided by the column total). For example, four mothers are in the shaded cell—they are younger mothers who had a repeat pregnancy. The expected count for that cell is three—if there were no relationship between age group and repeat pregnancy, we would expect three mothers to be in this cell, rather than four. Finally, 40% of the mothers who were 17 years or younger (compared to 25% who were 18 years or older) had an early subsequent pregnancy. Thus, the proportions are consistent with our hypothesis, but we need to test whether the differences could be spurious.

Panel B of Figure 5, which shows the results of statistical tests, has a lot of information that we did not discuss in the textbook. The most important information in this printout for our purposes is that the chi-square test ($\chi^2 = .714$) was not significant

---

[2]It is easy in SPSS to create new variables through various transformations. Some are described in Chapter 20.

A **Repeat pregnancy * mother's age group crosstabulation**

				Mother's age group		Total
				17 years or younger	18 years or older	
Repeat pregnancy	No	Count		6	15	21
		Expected count		7.0	14.0	21.0
		% within mother's age group		60.0%	75.0%	70.0%
	Yes	Count		4	5	9
		Expected count		3.0	6.0	9.0
		% within mother's age group		40.0%	25.0%	30.0%
Total		Count		10	20	30
		Expected count		10.0	20.0	30.0
		% within mother's age group		100.0%	100.0%	100.0%

B **Chi-square tests**

	Value	df	Asymp. Sig. (two-sided)	Exact. Sig. (two-sided)	Exact. Sig. (one-sided)
Pearson chi-square	.714[a]	1	.398		
Continuity correction[b]	.179	1	.673		
Likelihood ratio	.698	1	.403		
Fisher exact test				.431	.331
Linear-by-linear association	.690	1	.406		
N of valid cases	30				

a. one cell (25.0%) has expected count less than 5. The minimum expected count is 3.00

b. Computed only for a 2x2 table

FIGURE 5 SPSS printout for a crosstabulation: Maternal age group and repeat pregnancy status.

(p = .398). However, we can also see that the more appropriate test for these data is Fisher exact test. As indicated in the footnote, the expected frequency in one of the cells was less than 5. Thus, the results for Fisher's test should be reported: the two-tailed probability is .431. With such a small sample, we cannot conclude that maternal age is related to risk of a repeat pregnant.