Regression analysis on hypnosis data

Answer 1

Testing normal distribution of variables can be completed by relying on two methods – a statistical or a graphical one. Statistical, or objective, testing of normality is accomplished through the use of the Kolmogorov-Smirnov Test. However, such a statistical test is best to be used when sample sizes are either very small or very large. The current sample consisted of 165 participants, thus making it a moderate sample size for which the Kolmogorov-Smirnov Test may not be particularly informative. Thus, the choice was to rely on the graphical assessment of normality, which can be accomplished by looking at histograms and Q-Q plots. This subjective assessment of normality revealed that all variables, except Hypnotic Susceptibility and General Health, show a rather perfect bell-shaped curve on the histogram, with all data points falling closely to the "ideal" diagonal line on the Q-Q plot. Hypnotic Susceptibility seems to be slightly positively skewed, and General Health slightly negatively skewed. Yet, this skewedness and slight deviation from the diagonal line on the Q-Q plot seems just too low for these variables to affect the accuracy of the regression analysis.

Answer 2

The next step in the analysis is to test whether other assumptions of regression have been met. The first assumption states that there should be a linear relationship between independent and dependent variables, which can be checked by examining scatterplot graphs. This examination revealed that on each graph, showing an independent variable's correlation with Hypnotic Susceptibility, data

points make up a straight line, thus confirming that the assumption has been met. Moreover, the assumption which states that there should be no outliers in the data has been met only partially. This is revealed on the scatterplots for Hypnotic Susceptibility, Neuroticism, Extraversion, and General Health variables, where some data points fall far away from the graph's line. Finally, the assumption of homoscedasticity seems to be met, in so far on all graphs the variances along the line of best fit remain similar as the line progresses.

Answer 3

Characteristics of the sample were obtained by conducting descriptive and frequencies statistics. The sample consisted of a total of 165 participants, 82 (49.7%) of which were males and 83 (50.3%) of which were females. Their mean age was 39.76 (SD = 17.18), with age varying from 18 to 54. Male participants had a mean age of 39.43 (SD = 11.28) and female participants had a mean age of 40.12 (SD = 16.33).

Answer 4

To see if participants' levels of neuroticism affect the number of hours slept at night, the conducted analysis was Pearson correlation analysis. This analysis revealed a significant relationship between the two variables ($r(163) = -.31$, $p < .05$), indicating that higher levels of neuroticism relate to less hours of sleep. However, holding participants' general health constant within a partial correlation revealed a non-significant relationship between neuroticism and hours of sleep ($r(163) = -.10$, $p > .05$). What this implies is that the relationship between neuroticism and hours of sleep can actually be explained in terms of participants' general health. It is probable that it is actually general health

that affects hours of sleep, rather than neuroticism, and that neuroticism was previously found to relate to hours of sleep only because of the high overlap in the variance with participants' general health.

Answer 5

Regression analysis relied on the Entry method, thus entering all seven independent variables in the regression model simultaneously. This method was chosen because it was not certain which variable will act as a highest predictor of hypnotic susceptibility. Importantly, there was no multicollinearity in the data, as evident in the finding that none of the Tolerance statistics were below 0.1 and none of the VIF statistics were above 10. Durbin-Watson value was close enough to 2 to reveal that the assumption of independence of observations has been met as well.

The model reached significance ($F(6,158) = 17.93$, $p < .001$), explaining 43.4% of variance in hypnotic susceptibility scores. Predictor with highest strength of prediction was Hours of Sleep ($t(158) = -7.13$, $p < .001$), followed by Age ($t(158) = -5.99$, $p < .001$), and General Health ($t(158) = -6.13$, $p < .001$). Exact formulas for predicting the dependent variables can be made by using constants and *b* values (i.e., Hypnotic Susceptibility = 17.42-1.34(Hours of Sleep); Hypnotic Susceptibility = 17.42-.16(Age); Hypnotic Susceptibility = 17.42-.10(General Health)). Finally, the model's high accuracy of predictions is revealed in low standard errors for each predictor, and standard error of the model being lower than 2 (see Table 1).

Table 1. Correlations, Regression, and Descriptive statistics

	Correlations								Regression		
	Hypnot. suscept.	Age	Sex	IQ	Extraversion	Neuroticism	Hours of sleep	Gen. health	B	Beta	Sig
Age	-.29*	1	.03	-.09	-.06	-.20	-.48**	-.32*	-.16	-.63	.000
Sex	.10	.03	1	.21	.18	.22	-.33**	.53***	.06	.03	.358
IQ	.03	-.09	.21	1	.12	-.29*	.03	-.18	-.03	-.30	.489
Extraversion	.13	-.06	.18	.12	1	-.22	-.05	-.25	-.02	-.23	.153
Neuroticism	-.20	.22	-.29*	-.22	-.09	1	.08	-.11	-.03	-.20	.168
Hours of sleep	-.48**	-.33**	.03	-.05	.08	-.15	1		-1.34	-.64	.000
Gen. health	-.32*	.53***	-.18	-.25	-.11	.83**	-.41**	1	-.10	-.58	.000
	Descriptives										
MEAN	4.32	39.76	1.46	105.13	29.38	35.41	8.13	10.01	R Square = .548		
SD	1.12	17.18	.50	17.00	6.54	7.13	1.15	5.12	Adjusted R Square = .534		
									St. Error = 1.687		

*$p < .05$; **$p < .01$; ***$p < .001$