Homework 2

Q1. What is the mean father’s age?

28.9

Q2. What is the mean father’s age for low birthweight babies?

24.83

Q3. Is the father's age normally distributed? Justify your answer.

No. The father’s age has a positive skewness of 0.49 and the Shapiro test has a p-value of 0.0385 indicating there is a strong evidence against the father’s age being normal

Q4. If you apply the log transformation to the father's age, what is the mean score of the transformed variable?

3.337

Q5. Is the above mean score a good representation of the real value? Justify your answer.

No. the value 3.337 does not speak to the mean value for father’s age. It’s just a mean gotten as a result of transformation that we needed to do for our analysis.

Q6. Is the new variable (log transform of father’s age) normally distributed? Justify your answer.

Yes. the Shapiro test has a p-value of 0.1287 indicating there is a no evidence against the father’s age being normal

Q7. Is the variable “years father was in education” normally distributed?

No. The Shapiro test has a p-value of 0.00004844 indicating there is a strong evidence against the “years father was in education” being normal

Q8. Mentioning the null and alternative hypotheses, explain the above answer.

The null hypothesis states that father’s age is normal. The alternative hypothesis states that “years father was in education” is not normal. The p-value that we got indicates a strong evidence for the alternative hypothesis

Q9. What is the mean score for the variable “years father was in education” after you apply the Box-Cox transformation?

13.71

Q10. Is this new variable normally distributed? Explain.

No. the p-value is less than 0.05 indidcating that there is a strong evidence against the new variable being normally distributed.

Q11. What is the mean score for this new variable (B-C transformed fathers’ years in education) for mothers aged under 35?

13.55

Q12. Which test would you use to investigate the relationship between birth weight and father's age?

* Pearson product-moment correlation
* Spearman’s Rank order correlation
* Point-Biserial correlation
* Phi-Coefficient

Spearman’s Rank order correlation

Q13. Justify the above choice in terms of the distribution of data and the nature of the test.

The father’s age is not normally distributed.

Q14. What is the direction of that relationship?

Positive

Q15. What is the form of that relationship?

Non-normal

Q16. What is the degree of that relationship?

Correlation coefficient is 0.178

Q17. What test would you use to investigate the relationship between smoking and birth weight?

* Pearson product-moment correlation
* Spearman’s Rank order correlation
* Point-Biserial correlation
* Phi-Coefficient

Point-Biserial correlaton

Q18. Report on the above results including information about direction/form/degree of the relationship.

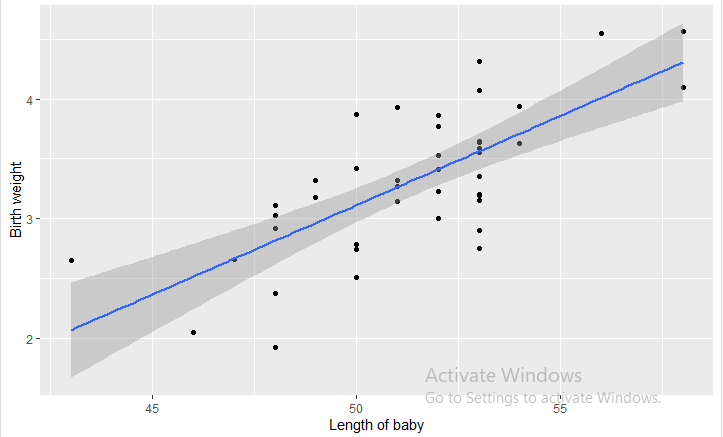
r= -3.14, p-value = 0.043, there is a statistically significant negative correlation

Q19. If you wanted to see the effect of the length of a baby on birthweight, what would your independent variable be?

* Length of baby
* Birthweight

Length of baby is the independent variable while birthweight is the dependent variable

Q20. In statistics, when creating a scatterplot, it is a common practice to put the independent variable on the x-axis and the dependent variable on the y-axis. With this in mind, create a scatterplot for the above case and provide the regression line. For homework submitted using MS Word, insert a picture of the scatterplot.



Q21. Is the relationship between the length of baby and birthweight linear?

* Yes
* No

Yes

Q22. Justify the above choice.

The plot above seem to indicate there is a linear relationship between the two variable

Q23. Is there any evidence to suggest that the birth weight, length of baby, and head circumference are related?

* Yes
* No

Yes

Q24. Justify the above choice.

There is a statistically significant relationship among all the 3 different pairs of the three variables

Q25. Describe the above relationship in your own words and provide evidence for your claims.

There is a significant relationship between head circumference and Birthweight with r = 0.685, p-value = 0.0001. There is a significant relationship between length of birth and Birthweight with r = 0.727, p-value = 0.0001. head circumference and length of birth with r = 0.563, p-value = 0.001