Lab 6: What Would Javi Do?
KEY

Javier E. Flores

April 1, 2019

Total Possible Points: 60

MRI and Alzheimer's

Q1) [6 pts] The Mini-Mental State Examination (MMSE) is a 30-point questionnaire used to measure cognitive impairment. This questionnaire is often used as a screener for dementia. At baseline, how different is the MMSE of subjects classified as demented from those classified as non-demented?

Data Scenario	Two-Sample Quantitative Data
Method chosen (include rationale)	Confidence interval for difference in means. Though the data are highly left skewed, the sample size in each group is large enough to apply the estimation procedure (think CLT).
Hypotheses (if applicable)	Not applicable
Statistic	95% CI: $\bar{x}_d - \bar{x}_{nd} \pm t_{crit,df=55}\sqrt{\frac{s_d^2}{n_d} + \frac{s_{nd}^2}{n_{nd}}} = (-4.95, -3.08)$
P-value (if applicable)	Not applicable
Brief Interpretation	We are 95% confident that the average MMSE of demented subjects at baseline is at least 3.08 points lower than non-demented subjects.

Q2) [6 pts] Typically, a MMSE score greater than or equal to 24 is indicative of normal cognition. At baseline, do subjects classified as demented fall below this threshold?

Data Scenario	One-Sample Quantitative
Method chosen (include rationale)	One sample t-test. Again, though the data are skewed, the sample is large enough to allow for application of this testing procedure.
Hypotheses (if applicable)	$H_0 : \mu_d \geq 24; \quad H_A : \mu_d < 24$
Statistic	$t_{test} = \frac{\bar{x}_d - \mu_0}{s_d/\sqrt{n_d}} = 2.58$
P-value (if applicable)	The left-tailed p-value is 0.99
Brief Interpretation	We fail to reject the null hypothesis that the mean MMSE among demented subjects at baseline is greater than or equal to 24.

Q3) [6 pts] Aside from the classifications of "Demented" or "Nondemented", the data contain a third category, "Converted", which are individuals who initially did not have dementia but later developed it. If we (unrealistically) assume that each of these patients had a perfect MMSE at baseline (i.e. 30), we would expect to see greater than a 6 point decrease in MMSE by their last visit (in order to be consistent with their classification of "converted"). For the "Converted" patients, is the MMSE at baseline greater than the last visit MMSE by more than 6 points?

Data Scenario	Paired Quantitative Data
Method chosen (include rationale)	Randomization test of differences. The data are measured on the same subjects but at two different time points. The sample size is relatively small (n = 14) and clearly non-normal.
Hypotheses (if applicable)	$H_0 : \mu_{diff} \leq 6; \quad H_A : \mu_{diff} > 6$
Statistic	Randomization distribution for a single mean assuming a null value of 6.
P-value (if applicable)	The right-tailed p-value is 1.
Brief Interpretation	We fail to reject the null hypothesis that the average difference in MMSE among "converted" subjects is less than or equal to 6.

Q4) [6 pts] Is the progression of dementia, as measured by the change in MMSE from baseline to the last visit, different between those initially classified as having dementia compared to those who "converted"?

Data Scenario	Two-Sample Quantitative Data
Method chosen (include rationale)	Randomization test for difference in means. Given the small sample size and non-normality in one of the groups, we should be conservative in our approach by using the randomization test. Otherwise, we run the risk of violating the normality assumption of the t-test procedure.
Hypotheses (if applicable)	$H_0 : \mu_{diff,dem} - \mu_{diff,con} = 0$; $H_A : \mu_{diff,dem} - \mu_{diff,con} \neq 0$
Statistic	Randomization distribution for the difference in means assuming a null value of 0.
P-value (if applicable)	The two-tailed p-value is 0.899
Brief Interpretation	We fail to reject the null hypothesis of no difference in progression between those who have converted and those who were initially classified as having dementia.

Q5) [6 pts] Normalized whole brain volume is widely considered to be a reliable method for quantifying neurodegeneration. Is the MMSE predictive of normalized whole brain volume?

Data Scenario	Quantitative Data (Explanatory and Response)
Method chosen (include rationale)	Regression. Our interest is whether we can use MMSE to predict the normalized whole brain value. We don't necessarily care about the strength of this association.
Hypotheses (if applicable)	$H_0 : \beta = 0$; $H_A : \beta \neq 0$
Statistic	b = 0.0037
P-value (if applicable)	The p-value is 0.
Brief Interpretation	We reject the null hypothesis that MMSE is not predictive of normalized whole brain value.

Police Killings

Q6) [6 pts] Recently, a jury in Pennsylvania acquitted a police officer who fatally shot an unarmed teenager in the back. How prevalent are police killings involving unarmed victims among all police killings?

Data Scenario	One-Sample Categorical Data
Method chosen (include rationale)	Confidence interval. Note that $np > 10$ and $n(1-p) > 10$.
Hypotheses (if applicable)	Not applicable
Statistic	95% CI: $\hat{p} \pm z_{crit}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0.18, 0.26)$
P-value (if applicable)	Not applicable
Brief Interpretation	We are 95% confident that the prevalence of police killings involving unarmed victims among all police killings is between 18% and 26%.

Q7) [6 pts] To describe the acquittal of the Pennsylvanian officer as well as the events surrounding the case, NPR wrote an article titled, "Jury Acquits White Former Police Officer In Fatal Shooting of Unarmed Black Teen". Is there a higher proportion of black unarmed deaths relative to white?

Data Scenario	Two-Sample Categorical Data
Method chosen (include rationale)	Two-Sample z-test. Both groups satisfy the conditions of $np > 10$ and $n(1-p) > 10$.
Hypotheses (if applicable)	$H_0 : p_b - p_w = 0; \quad H_A : p_b - p_w > 0$
Statistic	$z_{test} = \frac{(\hat{p}_b - \hat{p}_w) - 0}{\sqrt{\frac{\hat{p}_{pooled}(1-\hat{p}_{pooled})}{n_b} + \frac{\hat{p}_{pooled}(1-\hat{p}_{pooled})}{n_w}}} = 0.466$
P-value (if applicable)	The obtained p-value is 0.321.
Brief Interpretation	We fail to reject the null hypothesis of no difference in proportion of unarmed black and white police killing victims.

Q8) [6 pts] More generally, are the majority (i.e. more than half) of victims black?

Data Scenario	One-Sample Categorical Data
Method chosen (include rationale)	One-sample z-test. Note that $np > 10$ and $n(1-p) > 10$.
Hypotheses (if applicable)	$H_0 : p \leq 0.50; \quad H_A : p > 0.50$
Statistic	$z_{test} = \frac{\hat{p}-0.50}{\sqrt{\frac{p_0(1-p_0)}{n}}} = -9.12$
P-value (if applicable)	The p-value is 1.
Brief Interpretation	We fail to reject the null hypothesis that the proportion of black victims is less than or equal to 0.50.

Q9) [6 pts] The FiveThirtyEight article informed by these data states that "Police killings tend to take place in neighborhoods that are poorer and blacker than the U.S. as a whole." What proportion of killings are in black-majority neighborhoods whose census tracts are in the bottom 20% in terms of household income?

Data Scenario	One-Sample Categorical Data
Method chosen (include rationale)	Confidence interval. Again, np and $n(1-p)$ are both greater than 10.
Hypotheses (if applicable)	Not Applicable
Statistic	95% CI: $\hat{p} \pm z_{crit}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0.05, 0.10)$
P-value (if applicable)	Not Applicable
Brief Interpretation	We are 95% confident that the proportion of police killings in poor, majority-black neighborhoods is between 5% and 10%.

Q10) [6 pts] How different is this proportion from poor, majority hispanic communities?

Data Scenario	Two-Sample Categorical Data
Method chosen (include rationale)	Confidence interval. Here, both np and $n(1-p)$ are greater than 10.
Hypotheses (if applicable)	Not applicable.
Statistic	95% CI: $\hat{p}_b - \hat{p}_h \pm z_{crit}\sqrt{\frac{\hat{p}_b(1-\hat{p}_b)}{n} + \frac{\hat{p}_h(1-\hat{p}_h)}{n}} = (-0.02, 0.04)$
P-value (if applicable)	Not applicable.
Brief Interpretation	We are 95% confident that the difference in poor, majority black and poor, majority hispanic proportions is at least -0.02 and at most 0.04.