

# 2019Fall STAT5107 Assignment 1

Department of Statistics, The Chinese University of Hong Kong

Due 9:30pm, Thursday, September 26, 2019

1. Identify each variable as nominal, ordinal, or interval.
  - a. UK political party preference (Labour, Conservative, Social Democrat.)
  - b. Anxiety rating (none, mild, moderate, severe, very severe.)
  - c. Patient survival (in number of months.)
  - d. Clinic location (London, Boston, Madison, Rochester, Montreal.)
  - e. Response of tumor to chemotherapy (complete elimination, partial reduction, stable, growth progression.)
  - f. Favorite beverage (water, juice, milk, soft drink, beer, wine.)
  - g. Appraisal of company's inventory level (too low, about right, too high.)
2. For a binomial random variable  $B(n, \pi)$ , discuss whether it is easier to get a precise estimate of  $\pi$  when it is near 0 or 1 than when it is near  $\frac{1}{2}$ ? Please provide detailed arguments. (Hint: consider the variance of the estimate of  $\pi$ )
3. Consider the statement, "Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if she is married and does not want any more children." For the 1996 General Social Survey, conducted by the National Opinion Research Center(NORC), 842 replied "yes" and 982 replied "no". Let  $\pi$  denote the population proportion who would reply "yes". Find the P-value for testing  $H_0 : \pi = 0.5$  using the score test, and construct a 95% confidence interval for  $\pi$ .
4. In a crossover trial comparing a new drug to a standard,  $\pi$  denotes the probability that the new one is judged better. It is desired to estimate  $\pi$  and test  $H_0 : \pi = 0.5$  against  $H_a : \pi \neq 0.5$ . In 20 independent observations, the new drug is better each time.
  - a. Find and sketch the likelihood function. Give the maximum likelihood estimate of  $\pi$ .
  - b. Conduct a Wald test and construct a 95% Wald confidence interval for  $\pi$ .
  - c. Conduct a score test, reporting the P-value. Construct a 95% score confidence interval.
  - d. Conduct a likelihood-ratio test and construct a likelihood-based 95% confidence interval.
  - e. Suppose that researchers wanted a sufficiently large sample to estimate the probability of preferring the new drug to within 0.05, at confidence level 95%. If the true probability is 0.90, how large the sample size should be?
5. Let  $y_1, y_2, \dots, y_n$  be independent and identically distributed Poisson random variables with mean parameter  $\mu$ . Show that the maximum likelihood estimator of  $\mu$  is  $\hat{\mu} = \bar{y}$  and calculate the mean and variance of  $\hat{\mu}$ .

6. Table below contains Ladislaus von Bortkiewicz's data on deaths of soldiers in the Prussian army from kicks by army mules (Fisher 1934; Quine and Seneta 1987). The data refer to 10 army corps, each observed for 20 years. In 109 corps-years of exposure, there were no deaths, in 65 corps-years there was one death, and so on. Estimate the mean and test whether probabilities of occurrences in these five categories follow a Poisson distribution (truncated for 4 and above).

Number of Deaths	Number of Corps-Years
0	109
1	65
2	22
3	3
4	1
$\geq 5$	0