| | |
|---|---|
| COLUMBIA UNIVERSITY | Prof. Marco Avella |
| Statistical Inference | TA: George Chu, Joe Suk |
| Final | July 2, 2020 |

---

**Name:**

**UNI:**

---

You have 3 hours to complete this exam. Good luck!

## Question 1

Let $X_1 \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$. Find the maximum likelihood estimator $\mathbb{P}(X_1 = 0)$ and provide an asymptotic $1 - \alpha$ confidence interval for this quantity.

## Question 2

Suppose that a regular light bulb, a long-life light bulb, and an extra-long-life light bulb are being tested. The lifetime $X_1$ of the regular bulb has the exponential distribution with mean $\theta$, the lifetime $X_2$ of the long-life bulb has the exponential distribution with mean $2\theta$, and the lifetime $X_3$ of the extra-long-life bulb has the exponential distribution with mean $3\theta$.

1. Determine the MLE of $\theta$ based on $X_1, X_2, X_3$.
2. Let $\psi = 1/\theta$, and suppose that the prior distribution of $\psi$ is the gamma distribution with parameters $\alpha$ and $\beta$. Determine the posterior distribution of $\psi$ given $X_1, X_2, X_3$.

## Question 3

Let $N_1, \ldots, N_n \overset{iid}{\sim} Geometric(p)$, i.e. $\mathbb{P}(N_1 = k) = p(1-p)^{k-1}$ for $k = 1, 2, 3, \ldots$ and $p \in (0, 1)$. Note that this implies that $\mathbb{E}[N_1] = \frac{1}{p}$ and $\text{var}[N_1] = \frac{(1-p)}{p^2}$. Find the method of moments estimators of $p$ obtained using the first and second empirical moments respectively. Which of the two estimators would you expect to be a better one based on sufficiency? Show that it is also asymptotically efficient.

## Question 4

Let $X_1, \cdots, X_n$ be an iid random sample from the normal distribution with unknown mean $\mu$ and unknown standard deviation $\sigma$, and let $\hat{\mu}$ and $\hat{\sigma}$ denote their respective MLE. For the sample size $n = 17$, find a value of $k$ such that $\mathbb{P}(\hat{\mu} > \mu + k\hat{\sigma}) = 0.95$.

## Question 5

Suppose that a sequence of Bernoulli trials is to be carried out with an unknown probability $\theta$ of success on each trial, and the following hypotheses are to be tested

$$\begin{aligned} H_0: &\quad \theta = 0.1, \\ H_1: &\quad \theta = 0.2. \end{aligned}$$

Let $X$ denote the number of trials required to obtain a success, and suppose that $H_0$ is to be rejected if $X \leq 5$. Determine the power function of this test.

## Question 6

We ran a linear regression to a data on the relation between atmospheric pressure (inches of mercury) and the boiling point of water (degrees Fahrenheit) at different locations in the Alps. [1] A normal linear regression model with log pressure as the response led to the following `R` output:

```
Call:
lm(formula = log(pres) ~ bp, subset = -12)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0048082 -0.0014595  0.0004546  0.0020358  0.0031219

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.9517662  0.0231021   -41.2 5.16e-16 ***
bp           0.0205186  0.0001138   180.2  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.002616 on 14 degrees of freedom
Multiple R-squared:  0.9996,Adjusted R-squared:  0.9995
F-statistic: 3.249e+04 on 1 and 14 DF,  p-value: < 2.2e-16
```

Note also that the covariate is such that $\sum_{i=1}^n x_i = 3245.6$ and $\sum_{i=1}^n (x_i - \bar{x})^2 = 527.9$.

1. What test does p-value associated to `bp` correspond to? State explicitly the null and the alternative hypothesis. How would you interpret the result?
2. Compute a 0.95 confidence interval for the log pressure of an observation with bp$= 200$.
3. If we were to restrict ourselves to a model where the slope parameter is 0, what would be the MLE of the intercept $\beta_0$ and the variance $\sigma^2$?
4. Give the likelihood ratio statistic for testing $H_0 : \beta_1 = 0$.

---

[1] The data is due to James Forbes, a Scottish physicist from the 19th century, who sought a formula for predicting pressure $y$ from boiling point $x$.