Math 472: Homework 05

Due Friday Feb 27

Problem 1 (Exercise 8.31)**.** In a study to compare the perceived effects of two pain relievers, 200 randomly selected adults were given the first pain reliever, and 93% indicated appreciable pain relief. Of the 450 individuals given the other pain reliever, 96% indicated experiencing appreciable relief.

(a) Give an estimate for the difference in the proportions of all adults who would indicate perceived pain relief after taking the two pain relievers. Provide a bound on the error of estimation.

(b) Based on your answer to part (a), is there evidence that proportions experiencing relief differ for those who take the two pain relievers? Why?

Problem 2 (Exercise 8.17)**.** If Y has a binomial distribution with parameter n and p, the $\widehat{p} = \frac{Y}{n}$ is an unbiased estimator of p. Another estimator is $\widehat{q} = \frac{Y+1}{n+2}$.

(a) What is the bias of $\widehat{q}$?

(b) Derive $\mathtt{MSE}(\widehat{p})$ and $\mathrm{MSE}(\widehat{q})$.

(c) For what values of p is $\mathrm{MSE}(\widehat{p}) < \mathrm{MSE}(\widehat{q})$?

Problem 3 (Exercise 8.39)**.** Suppose the random variable Y has a gamma distribution with shape parameter $\alpha = 2$ and unknown scale parameter β (see section 4.6 in textbook). It can be shown that $2Y/\beta$ has a chi-squared distribution with 4 degrees of freedom (you don't have to show this).

(a) Using $2Y/\beta$ as a pivotal quantity, derive a 90% confidence interval of β.

(b) Using R, test your confidence interval by generating 10,000 samples $Y_1, \ldots, Y_{10000}$ (each with a randomly chosen numerical value for β) and counting the proportion of times that your confidence interval contains the true value of β. (It should be close to 90%).

Problem 4 (Exercise 8.40)**.** Suppose that the random variable Y is an observation from a normal distribution with unknown mean μ and variance 1.

(a) Find a 95% confidence interval for μ.

(b) Find a 95% upper confidence interval for μ.

(c) Find a 95% lower confidence interval for μ.

Problem 5 (Excercise 8.41)**.** Suppose that Y is normally distributed with mean $\mu = 0$ and unknown variance σ^2. Then Y^2/σ^2 has a chi-squared distribution with 1 degree of freedom. Use the pivotal quantity Y^2/σ^2 to find

(a) A 95% confidence interval for σ^2.

(b) A 95% upper confidence interval for σ^2.

(c) A 95% lower confidence interval for σ^2.

Problem 6 (Part of exercise 6.18)**.** The Pareto distribution is a "power-law" distribution used for modeling things like income distributions and earthquake magnitudes. Specifically, given a random variable X, we say that distribution of X is a member of the ***Pareto family of distributions*** if its distribution function is

$$F_X(x) = \mathbb{P}\left[X \leq x\right] = \begin{cases} 1 - \left(\frac{\beta}{x}\right)^\alpha & : \quad x \geq \beta \\ 0 & : \quad x < \beta \end{cases} \tag{1}$$

for some positive parameters α and β.

(a) Big-O notation: Let f and g be real-valued functions whose domains include the positive real numbers. We write $f(x) = O(g(x))$ as $x \to \infty$ if there exist constants $M, k > 0$ such that

$$\frac{|f(x)|}{|g(x)|} \leq M$$

for all $x > k$. (This is read as "f is big-O of g".)

Show that $\mathbb{P}[X \geq x] = O(\frac{1}{x^\alpha})$ as $x \to \infty$. (In words, this says that the upper tail of the probability distribution of X decays like $\frac{1}{x^\alpha}$.)

(b) Derive the density function of X. Make sure to specify the function on all parts of its domain: both when $x \geq \beta$ and when $x < \beta$.

(c) Show that the improper integral $\int_1^\infty \frac{1}{x^\delta} dx$ diverges to $+\infty$ if $\delta \leq 1$, but converges to $\frac{1}{\delta - 1}$ if $\delta > 1$.

(d) What is $\mathbb{E}[X]$? Consider the cases $\alpha \leq 1$ and $\alpha > 1$ separately.

Problem 7 (Part of exercise 9.28)**.** Let $X_1, \ldots, X_n$ be a random sample from the distribution F_X (defined in Eq. (1)), and let T be the statistic $T(X_1, \ldots, X_n) = \min(X_1, \ldots, X_n)$.

(a) Show that the sampling distribution of T is

$$F_T(t) = \begin{cases} 1 - \left(\frac{\beta}{t}\right)^{\alpha n} & : \quad t \geq \beta \\ 0 & : \quad t < \beta \end{cases} . \tag{2}$$

(*Hint:* For every real number t, we have $[\min(X_1, X_2, \ldots, X_n) > t] = [X_1 > t, X_2 > t, \ldots, X_n > t]$).

(b) Derive the density function of T from Eq. (2). Make sure to specify the function on all parts of its domain: both when $t \geq \beta$ and when $t < \beta$.

(c) Qualitatively describe what happens to the graph of f_T as n becomes very large. Sketch this situation, and interpret your picture: is T likely to be close to β or far from β when n is large?

Problem 8 (Exercise 8.15)**.** Let $X_1, \ldots, X_n$ be a random sample from the Pareto distribution with probability density function

$$f_X(x) = \begin{cases} 3\beta^3 x^{-4} & : \quad x \geq \beta \\ 0 & : \quad x < \beta \end{cases} ,$$

where $\beta > 0$ is unknown. (So X is a Pareto distribution with $\alpha = 3$). We shall consider the estimator $\widehat{\beta} = T(X_1, \ldots, X_n) = \min(X_1, \ldots, X_n)$.

(a) What is the bias of $\widehat{\beta}$?

(b) What is $\mathrm{MSE}(\widehat{\beta})$?