Exam #1

Economics 435

Spring 2004

Note: This exam has been altered slightly to include the clarifications and corrections made during class.

1 Airplanes

The idea for this problem comes from a supposedly true story about the Royal Air Force during World War II. Halfway through the war, RAF engineers figured out that they could add additional armor reinforcement to their bombers which would protect the plane against enemy fire. However, they could not reinforce the entire plane because then the plane would be too heavy to fly. The way they decided which parts to reinforce was as follows: they looked at planes returning from bombing runs and counted the number of bullet marks in each section of the plane. They then reinforced those sections of the plane with the most bullet marks, based on the reasoning that those were the sections that seem to have been shot at the most. This was not necessarily a wise decision.

We will simplify the problem by supposing the plane can be split into two halves (call them left and right, or L and R). On a given bombing run the plane will either be hit on the left (with probability p_L), be hit on the right (with probability p_R), or will escape without damage (with probability $1 - p_L - p_R$). A hit on the left side will cause the plane to crash with probability q_L and a hit on the right side will cause the plane to crash with probability q_R. The plane will not crash if it is not hit. The RAF observes the fraction of returning (i.e., not crashed) planes that are hit on the left ($\hat{p}_L$), and the fraction that are hit on the right ($\hat{p}_R$). Suppose that more of the returning planes have damage on the left (i.e., that $\hat{p}_L > \hat{p}_R$). It also knows the fraction of planes that crash $\hat{c}$, though it does not know what caused those crashes. Suppose that the number of planes and bombing runs are sufficiently large that each of these proportions is equal to the underlying probability.

a) The RAF has the goal of minimizing the probability of crashing. What is that probability, in terms of p_L, p_R, q_L, and q_R?

b) What are the expected value of $\hat{p}_L$, $\hat{p}_R$, and $\hat{c}$ in terms of p_L, p_R, q_L, and q_R?

c) Now we consider the problem faced by the RAF. Suppose you could reinforce either the left or the right, but not both. Also suppose that the reinforcement makes the plane invulnerable on that side (for example, if you reinforce the right then the probability of crashing conditional on being hit on the right is zero). What would the probability of crashing be if the left were reinforced, in terms of p_L, p_R, q_L, and q_R? What would the probability of crashing be if the right were reinforced?

d) Does the RAF have enough information to make a decision? Explain.

e) Suppose that the RAF believes that, without reinforcements, both the left and right sections are equally well-protected, i.e. that $q_L = q_R$. Which side should they reinforce? Use explicit probability calculations to support your answer.

f) Suppose that the RAF believes that both sides are equally likely to be hit by enemy fire, i.e., that $p_L = p_R$, Which side should they reinforce? Use explicit probability calculations to support your answer.

g) Suppose you don't believe either of these assumptions, but the RAF is willing to let you run a 1-month test program in which 100 planes are reinforced (you get to pick which side each plane is reinforced on), and used in service. How do you organize your test program, and what do you do with the results?

2 Income and life expectancy

Suppose we are interested in the relationship between a country's income level (GDP/capita) and the life expectancy of its citizens. We have a cross-country data set and estimate the following model

$$LifeExp = \beta_0 + \beta_1 GDP + \beta_2 GDP^2 + u$$

where $LifeExp$ is the average life expectancy in years, and GDP is 2002 GDP/capita in thousands of US dollars. The regression results (number of observations = 170) are as follows:

Coefficient	Estimate	Standard Error
$\hat{\beta}_0$	54.564	1.0729
$\hat{\beta}_1$	1.9375	0.1923
$\hat{\beta}_2$	-0.0353	0.0058

a) Perform an asymptotic hypothesis test at the 5% level of significance for the null hypothesis that $LifeExp$ is linearly related to the country's GDP per capita, i.e., that $\beta_2 = 0$. State the null and alternative, write down the formula for the test statistic you are using, and its asymptotic distribution under the null, calculate the test statistic, state the rule for rejecting/accepting the null, and report your accept/reject decision.

b) Calculate the estimated marginal benefit in terms of additional years of life expectancy from a increase in per capita GDP of ΔGDP units, (where ΔGDP is small) as a function of current GDP. Use calculus.

c) Sierra Leone has the lowest per capita GDP in the sample, at \$490 per year, and one of the lowest life expectancies, at 39 years. By how much does this model predict the life expectancy in Sierra Leone would increase if its per capita GDP reached the median value in our sample ($4,733)? Do not use calculus.

3 The mode

Sometimes[1] we are interested in estimating the *mode* of a discrete probability distribution, which is just the most likely value. Specifically, let X be a random variable which takes on values from a finite set K, and let $f(x) \equiv \Pr(X = x)$ be the PDF of X. The mode (M) of this distribution[2] is defined as:

$$M \equiv \arg \max_{z \in K} \Pr(X = z)$$

Suppose you have a random sample of size N on X: $\{x_1, x_2, \ldots, x_N\}$.

a) Use the analog principle to suggest an estimator for M. Note that $\Pr(X = z) = E[I(x = z)]$, where $I(.)$ is the indicator function (returns a value of 1 for a true statement, 0 for a false statement).

b) Is your estimator unbiased? Construct a simple example (Hint:try one with $N = 1$ and $K = \{0, 1\}$) to support your answer.

c) Is your estimator consistent? Provide a proof to support your answer.

d) For extra points: Is your estimator likely to be asymptotically normal? Explain.

[1] An example, if you don't believe me: Criminologists and police departments are often interested in when and where crimes take place. They will frequently want to know the locations where the most crimes take place (so they can put more officers there); it makes no sense to look for the "average" location of crimes.

[2] The mode of a distribution is not unique for all probability distributions, but for this problem we will assume we have a distribution with a single mode.