Inference and Regression

Midterm Examination, 2016

- Instructions
  - Please write your name at the top of this page.
  - Please answer all questions on this question book. Do not turn in a blue book.
  - Please do not separate the pages of this exam booklet.
  - There are 11 questions in this exam. Questions 1 (25 points), 2 (20 points) and 3 (15 points) are mandatory. Please answer 6 of the remaining 8 questions 4 – 11. All are worth 15 points, so your total points scored for the exam will be 25 + 20 + 15 + 6(15) = 150.
  - Where a computation is required to answer a question, please show your work.
    (I cannot give partial credit for an incorrect numerical answer unless the work provided shows a partially correct computation.)

- Introduction

Several of the questions below are based on the Weibull distribution for a continuous, nonnegative random variable. The Weibull distribution depends on 2 parameters. I have fixed one of them at 2. The density of the Weibull(k=2) random variable, which depends on one parameter, \( \lambda \), is

\[
f(x | \lambda) = \frac{2x}{\lambda^2} \exp\left(-\frac{x^2}{\lambda^2}\right), \quad x \geq 0, \quad \lambda > 0.
\]

Several moments of this random variable are

\[
E[x] = \lambda \Gamma(1+1/2) = \lambda (1/2) \Gamma(1/2) = \lambda \times 1/2 \times \sqrt{\pi}
\]

\[
E[x^2] = \lambda^2
\]

Variance(x) = \( E[x^2] - (E[x])^2 \) = \( \lambda^2 - \lambda^2 \pi/4 \) = \( \lambda^2 (1 - \pi/4) \)

Standard deviation(x) = \( \lambda \sqrt{1 - \pi/4} \)
Elizabeth has been selling cupcakes out of her shop at 30th and M in Washington for many years. She knows by now that daily sales are normally distributed with mean $\mu$ and $\sigma = 240$. Based on previous experience, $\mu = 1000$. But, Allison opened a shop down the street (at Wisconsin Avenue) and it looks like she took some business away. If demand has really fallen off significantly, Elizabeth will have to lay off at least one baker (Julie). This is very costly. In order to find out, she will carry out a test. The strategy is as follows: Watch demand for 64 days and compute the mean demand, $\bar{x}$. The standard deviation, $\sigma$, is known to be 240. The rejection region is $\bar{x} \leq 950$. If the sample data fall in the rejection region, Baker Julie loses her job. If the data do not fall in the rejection region, Julie stays and everyone is happy.

a. What is the null hypothesis for this test? What is the alternative?
b. What is the probability of a type 1 error for the strategy of rejecting $H_0$ if $\bar{x} \leq 950$?
c. What is the probability of a type 2 error for this strategy if $\mu = 975$?
d. What is the power of the test using strategy if $\mu = 990$?
e. To be conservative, Elizabeth will repeat the experiment three more times (i.e., watch demand for 64 days, three times and calculate $\bar{x}$ each time). What is the probability that she will reject the null hypothesis in part a. at least twice based on these three experiments?

a. $H_0: \mu \geq 1000$
b. $\text{Prob}(x-\bar{<}950|\mu=1000,\sigma=240) = \text{Prob}(z<(950-1000)/(240/\sqrt{64}))$
   $= \text{Prob}(z<-50/30) = .04779$
c. $\text{Prob}(\bar{x} > 950 | \mu = 975, \sigma = 240) = \text{Prob}(x > (950 - 975)/(240/\sqrt{64}))$
   $= \text{Prob}(z > -25/30) = .7977$
d. $\text{Prob}(\bar{x} < 950|\mu = 990, \sigma = 240) = \text{Prob}(z < (950 - 990)/(240/\sqrt{64}))$
   $= \text{Prob}(z < -40/30) = .0912$
e. From part b, prob reject on each try is .04779. Prob reject 2 or 3 times is
   $3C2 \times .04779^2 (1 - .04779) + 3C3 \times .04779^3 = .00663$

[20] 2. Derive the maximum likelihood estimator of $\lambda$ for the Weibull distribution discussed in the introduction based on a sample of N observations, $x_1, \ldots, x_N$. Find the variance of the maximum likelihood estimator.

$$f(x|\lambda) = \frac{2x}{\lambda^2} \exp\left(-\frac{x^2}{\lambda^2}\right), \ x \geq 0, \ \lambda > 0.$$ 

$$\ln L = \sum_i \left(\ln 2 + \ln x - 2\ln \lambda - \frac{x^2}{\lambda^2}\right)$$

$$d\ln L/d\lambda = -2N/\lambda + 2\Sigma x^2/\lambda^3$$

set equal to zero, $2N\lambda^2 = 2\Sigma x^2$ so $\lambda^2 = \text{sqr}[(1/N)\Sigma x^2]$. Note, this makes sense. $(1/N)\Sigma x^2$ estimates $E[x^2] = \lambda^2$. MLEs are invariant, so to estimate the square root of the expectation, we use the square root of the estimator.

To get the variance,

$$d^2\ln L/d\lambda^2 = 2N/\lambda^2 - 6\Sigma x^2/\lambda^4.$$ Get the expected value. $E[x^2] = \lambda^2$ so the expected value of the second derivative is $2N/\lambda^2 - 6N\lambda^2/\lambda^4 = -4N/\lambda^2$. The negative of the reciprocal of this gives the variance, which would be $\lambda^2/(4N)$. 


This question is based on the Weibull distribution discussed in the introduction. Based on the characteristics given for the distribution:

1. Obtain two different method of moments estimators of \( \lambda \) based on \( E[x] \) and \( E[x^2] \).
2. What is the variance of the estimator based on \( E[x] \)? (Hint, this is a standard result.)

1. \( E[x] = \sqrt{\pi} \lambda^2 / 2 \), so one estimator would be \( 2 / \sqrt{\pi} \times x-bar \). Since \( E[x^2] = \lambda^2 \), the second estimator is the MLE from the previous question.
2. \( \text{Var}[x-bar] = \text{Var}[x] / N = (1/N) \{E[x^2] - (E[x])^2\} = (1/N)\lambda^2 (1 - \pi/4) / N \). Then, \( \text{Var}[2/\sqrt{\pi} \times x-bar] = (4/\pi) \times \text{Var}[x-bar] = (4/\pi - 1) \lambda^2 / N \).

A Lucky People In Touch Magazine survey asked people about whether they were addicted to “Keeping Up with the Kardashians.” Their survey results suggested the following population results. Denote actually addicted by \( A \). Denote by \( R \) someone who admits they are addicted when asked.

- \( P(A) = .20 \) 20% of the population is actually addicted.
- \( P(R|A) = .40 \) 40% of addicts will admit they are addicted.
- \( P(\text{notR}|\text{notA}) = .95 \) 95% of people who are not addicted say they are not. The other 5% are having fun with the interviewer.

a. What is the probability that someone who says they are addicted actually are addicted.

b. Suppose everyone who was not addicted said they were not addicted when asked. That is, suppose \( P(\text{notR}|\text{notA}) = 1.00 \). What would the answer to part a. be in this case?

a. \( P(A|R) = P(R|A)P(A) / [P(R|A)P(A) + P(R|\text{notA})P(\text{notA})] \).
   We need \( P(R|\text{notA}) = 1 - P(\text{notR}|\text{notA}) = 1 - .95 = .05 \).
   \( P(\text{notA}) = 1 - .2 = .8 \).
   Result = .4(.2) / (.4(.2) + .05(.8)) = .08 / (.08 + .02) = .8

b. If \( P(\text{notR}|\text{notA}) = 1 \), \( P(R|\text{notA}) = 0 \) and \( P(A|R) = 1 \). The second term in the denominator underlined above would be 0.

FeelSmart pharmaceuticals is testing a new drug that they are convinced make cranky people feel (and act) much smarter. In order to market the drug, of course, they carried out a clinical trial. In a sample of 600 people, their clinical trial results were as follows:

<table>
<thead>
<tr>
<th>Response to the Treatment</th>
<th>None</th>
<th>Smarter</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>84</td>
<td>163</td>
<td>247</td>
</tr>
<tr>
<td>The Actual Drug</td>
<td>42</td>
<td>311</td>
<td>353</td>
</tr>
<tr>
<td>Total</td>
<td>126</td>
<td>474</td>
<td>600</td>
</tr>
</tbody>
</table>

a. Are these data consistent with the claim that the drug makes people feel smarter?

The question is ambiguous. One might be tempted to carry out a contingency table test. But, the question does not ask if the smarter and drug are independent, it specifically asks if people who got the drug feel smarter. Strictly, one might use a test of means. Is the 311/353 = .881 significantly greater than .5? The standard test statistic would be \( (.881 - .5)/\sqrt{.881 \times .119/353} = .381/0.17 = 22.1 \), so yes. To carry out the contingency test, instead, do the standard calculations. the actual proportions are .140,.272 / .070,.518. The fitted proportions are .086,.325/.124,.465. The chi squared is \( (.140-.086)^2/0.086 + (.272-.325)^2/0.325 + (.070-.124)^2/0.124 + (.518-.465)^2/0.465 \) * 600 = 43.264

The critical value for a 2x2 table is 3.84, so the independence hypothesis is rejected.
b. Some food for thought: Suppose instead of asking people the simple question of whether they are smarter or not, the analysts actually tested how smart their subjects were. They did two tests, one on March 20, 2014 and one on May 10, 2014. For the people who got the real drug, their average smart score rose from 52 to 83. For those who got the placebo, their average score rose from 48 to 74. What should the researchers conclude? Explain.

You would want to examine the difference in the changes, Drug(83 – 52) – Placebo(74 – 48) = 31 – 26 = 5. Not much. Now, is the difference significant? You need some data on the individual scores to form a variance.

[15] 6. The CDF for the Weibull distribution discussed in the introduction is \( F(x) = 1 - \exp(-\frac{x^2}{\lambda^2}) \).
For the moment assume you know \( \lambda = 3 \). Assuming you can use a random number generator to produce a sample of values from the uniform[0,1] population, how could you use this sample to produce a sample of draws from the Weibull population.

Treat a random draw, U as a draw on the CDF. The \( n \ln(1 - U) \) is a draw on \(-\frac{x^2}{\lambda^2}\). Since \( \lambda = 3 \), \( 9 \times \ln(1 - U) \) is a draw on \(-x^2\). So, \( x = -\sqrt{9 \times \ln(1 - U)} \).

[15] 7. This question is based on the Weibull distribution discussed in the introduction. The entropy of the random variable is

\[
\text{Entropy}(x) = 1.2886 + \ln(\frac{\lambda}{2})
\]

Suppose you have computed your maximum likelihood estimator of \( \lambda \) and it is 2.0. You have also estimated the asymptotic variance of your estimator and your estimated asymptotic variance is 0.7. Estimate \( \text{Entropy}(x) \) and obtain an estimator of the asymptotic variance of this estimator of \( \text{Entropy}(x) \). (Note, you do not need to know what the entropy is to answer this question. The question is about the variance of a function of the MLE.)

Use the delta method. The estimate of \( \lambda \) is 2.0. The estimated asymptotic is 0.7. The derivative you need is \( d\text{Entropy}/d\lambda = 1/\lambda \). So, the estimate of the asymptotic variance is \((1/\lambda)^2 \times 0.7 = 1/\lambda \times 0.7 = 0.175\).

[15] 8. Suppose \( x_1 \) and \( x_2 \) are independent normally distributed variables with means 3 and 4, respectively and standard deviations 1 and 5, respectively. \( z \) is distributed as Poisson with parameter 6 and is independent of \( x_1 \) and \( x_2 \). Two new random variables are formed as \( y_1 = x_1 + z \) and \( y_2 = x_2 + z \). What are the variances of \( y_1 \) and \( y_2 \)? What is the correlation between \( y_1 \) and \( y_2 \)? What is the correlation between \( y_1 \) and \( z \)?

\[
\begin{align*}
\text{Var}[y_1] &= \text{Var}[x_1] + \text{Var}[z] = 1^2 + 6 = 7 \\
\text{Cov}[y_1,y_2] &= \text{Var}[z] = 6. \quad \text{Var}[y_1] = 7, \text{Var}[y_2] = 25+6 = 31. \quad \text{Corr}(y_1,y_2) = 6/\text{sqr}[7\times31] = .407 \\
\text{Cov}[y_1,z] &= \text{Var}[z] = 6. \quad \text{Corr}[y_1,z] = 6/\text{sqr}[6\times7] = .926.
\end{align*}
\]

[15] 9. For the Weibull random variable discussed in the introduction, demonstrate whether or not the distribution is an exponential family. If it is an exponential family, what is the sufficient statistic? Explain.
From part 2, \( \ln L = \Sigma_i (\ln 2 + \ln x - 2\ln \lambda - x^2/\lambda^2) \). This is of the form constant + c(data) + d(parameters) + e(data)^f(parameters), which makes it an exponential family. The sufficient statistic is \( \Sigma_i x_i^2 \). Note in part 2, \( \lambda^2 = \text{sqr}[(1/N)\Sigma_i x_i^2] \). The MLE will be a function of the sufficient statistic.
a. Explain the difference between consistency and unbiasedness.

Unbiasedness relates to the expected value of an estimator, but makes no statement about its variance or distribution. Consistency implies that the variance of the estimator goes to zero, or that the distribution collapses to a point as \( n \) gets large. Unbiasedness does not relate to the sample size. Consistency implies an improvement of an estimator as the sample size increases.

b. Explain the difference between the law of large numbers and the central limit theorem.

Law of large numbers states that statistics converge in probability to their population counterparts. It says nothing about the distribution of the statistic. The CLT states that a statistic (more precisely, \( \sqrt{n}[\text{statistic} - \text{mean}] \)) has a distribution that converges to the normal. CLT relates to the distribution of a statistic. LLN relates to the probability limit of a statistic.

11. Referring to the Weibull random variable discussed in the introduction, suppose \( z = x^2 \). What is the density of \( z \)?

\[
f(x|\lambda) = \frac{2x}{\lambda^2} \exp\left(-\frac{x^2}{\lambda^2}\right), \ x \geq 0, \ \lambda > 0.
\]

If \( z = x^2 \) then \( x = \sqrt{z} \) and \( dx/dz = 1/[2\sqrt{z}] \). Make the substitution of \( z \) for \( x \) and multiply by the Jacobian.

\[
f(z|\lambda) = 2\sqrt{z}/\lambda^2 \times \exp\left(-\frac{\sqrt{z}^2}{\lambda^2}\right) \times 1/[2\sqrt{z}].
\]

The 2\sqrt{z}'s cancel. \( \sqrt{z}^2 = 2 \). So,

\[
f(z|\lambda) = 1/(\lambda^2) \exp\left[-z/\lambda^2\right].
\]

This is an exponential distribution with parameter \( \gamma = \lambda^2 \).