# STA 431s17 Assignment Five[1]

This assignment is on large-sample likelihood ratio tests, and the first part of measurement error. The material on likelihood ratio tests is in lecture slide set 8. See also Section A.6.5 in Appendix A, especially pages 176-178. The material on measurement error is in lecture slide set 9. See also Section 0.7 in Chapter Zero, pages 33-38.

1. Let $Y_1, \ldots, Y_n$ be a random sample from a distribution with density $f(y) = \frac{1}{\theta} e^{-\frac{y}{\theta}}$ for $y > 0$, where the parameter $\theta > 0$. We are interested in testing $H_0 : \theta = \theta_0$. This is an exponential distribution and the MLE is $\overline{Y}$. You don't have to re-derive it.
   (a) What is the parameter space $\Theta$?
   (b) What is the restricted parameter space $\Theta_0$?
   (c) Calculate a formula for the log likelihood evaluated at the unrestricted MLE. Simplify.
   (d) Derive a general expression for the large-sample likelihood ratio statistic $G^2$.
   (e) What is the distribution of the test statistic under the null hypothesis? Don't forget the degrees of freedom.
   (f) A sample of size $n = 100$ yields $\overline{Y} = 1.37$ and $S^2 = 1.42$. One of these quantities is unnecessary and just provided to irritate you. Well, actually it's a mild substitute for reality, which always provides you with a huge pile of information you don't need. Anyway, we want to test $H_0 : \theta = 1$. You can do this with a calculator. When I did it a long time ago I got $G^2 = 11.038$.
   (g) What is the critical value at $\alpha = 0.05$? The answer is a number from the formula sheet.
   (h) Do you reject $H_0$ at $\alpha = 0.05$? Answer Yes or No.
   (i) Are the results statistically significant? Answer Yes or No.
   (j) Is there evidence that $\theta \neq 1$? Answer Yes or No.
   (k) Choose one of these conclusions: $\theta < 1$, $\theta = 1$ or $\theta > 1$.

2. The label on the peanut butter jar says peanuts, partially hydrogenated peanut oil, salt and sugar. But we all know there is other stuff in there too. In the United States, the Food and Drug administration requires that a shipment of peanut butter be rejected if it contains an average of more than 8 rat hairs per pound (well, I'm not sure if it's exactly 8, but let's pretend). There is very good reason to assume that the number of rat hairs per pound has a Poisson distribution with mean $\lambda$, because it's easy to justify a Poisson process model for how the hairs get into the jars. The Poisson probability mass function is $p(y) = \frac{e^{-\lambda}\lambda^y}{y!}$, where $\lambda > 0$. The MLE is the sample mean; you don't have to re-derive it. We will test $H_0 : \lambda = \lambda_0$.
   (a) What is the parameter space $\Theta$?
   (b) What is the restricted parameter space $\Theta_0$?
   (c) Calculate a formula for the log likelihood evaluated at the unrestricted MLE. Simplify.

(d) Derive a general expression for the large-sample likelihood ratio statistic $G^2$.

(e) What is the distribution of the test statistic under the null hypothesis? Don't forget the degrees of freedom.

(f) We sample 100 1-pound jars, and observe a sample mean of $\overline{Y} = 8.57$ rat hairs. Should we reject the shipment? We want to test $H_0 : \lambda = 8$. What is the value of $G^2$? The answer is a number. You can do this with a calculator. When I did it a long time ago I got $G^2 = 3.97$.

(g) Do you reject $H_0$ at $\alpha = 0.05$? Answer Yes or No.

(h) Do you reject the shipment of peanut butter? Answer Yes or No.

3. Let $X_1, \ldots, X_n \stackrel{i.i.d}{\sim} N(\mu, \sigma^2)$. We want to test $H_0 : \sigma^2 = \sigma_0^2$ versus $\sigma^2 \neq \sigma_0^2$. You should know the unrestricted MLE for the normal model; you don't have to re-derive it.

(a) What is the parameter space $\Theta$?

(b) What is the restricted parameter space $\Theta_0$?

(c) Calculate a formula for the log likelihood evaluated at the unrestricted MLE. Simplify.

(d) Derive a general expression for the large-sample likelihood ratio statistic $G^2$.

(e) What is the distribution of the test statistic under the null hypothesis? Don't forget the degrees of freedom.

(f) A random sample of size $n = 50$ yields $\overline{X} = 9.91$ and $\widehat{\sigma}^2 = 0.92$. Calculate the test statistic; the answer is a number.

(g) What is the critical value of the test statistic at $\alpha = 0.05$?

(h) Do you reject $H_0 : \sigma^2 = 1$ at $\alpha = 0.05$? Answer Yes or No.

(i) Are the results statistically significant? Answer Yes or No.

(j) What, if anything, do you conclude?

4. You might want to look again at the coffee taste test example from lecture Unit 5 (Estimation) before starting this question. An email spam company designs $k$ different emails, and randomly assigns email addresses (from a huge list they bought somewhere) to receive the different email messages. So, this is a true experiment, in which the message a person receives is the experimental treatment. $n_1$ email addresses receive message 1, $n_2$ email addresses receive message 2, ..., and $n_k$ email addresses receive message $k$.

The response variable is whether the recipient clicks on the link in the email message: $Y_{ij} = 1$ if recipient $i$ in Treatment $j$ clicks on the link, and zero otherwise. According to our model, all these observations are independent, with $P(Y_{ij}) = \theta_j$ for $i = 1, \ldots, n_j$ and $j = 1, \ldots, k$. We want to know if there are any differences in the effectiveness of the treatments.

(a) What is the parameter space $\Theta$?

(b) What is the restricted parameter space $\Theta_0$?

(c) What is $\Theta_0$?

(d) Write the log likelihood function and simplify.

(e) What is $\widehat{\boldsymbol{\theta}}$? If you think about it you can write down the answer without doing any work.

(f) What is $\widehat{\boldsymbol{\theta}}_0$? If you think about it you can write down the answer without doing any work.

(g) Write down and simplify a general expression for the large-sample likelihood ratio statistic $G^2$. What are the degrees of freedom?

(h) Comparing three spam messages with $n_1 = n_2 = n_3 = 1,000$, the company obtains $\overline{Y}_1 = 0.044$, $\overline{Y}_2 = 0.050$ and $\overline{Y}_3 = 0.061$.

(i) What is the test statistic $G^2$? The answer is a number.

(j) What is the critical value at $\alpha = 0.05$? The answer is a number from the formula sheet.

(k) Do you reject $H_0$ at $\alpha = 0.05$? Answer Yes or No.

(l) Are the results statistically significant? Answer Yes or No.

(m) Is there evidence that the messages differ in their effectiveness? Answer Yes or No.

5. You may think of this as a continuation of Question 5 of Assignment 2. Let $Y_i = \beta x_i + \epsilon_i$ for $i = 1, \ldots, n$, where $\epsilon_1, \ldots, \epsilon_n$ are a random sample from a normal distribution with expected value zero and variance $\sigma^2$. The parameters $\beta$ and $\sigma^2$ are unknown constants. The numbers $x_1, \ldots, x_n$ are known, observed constants.

   (a) What is the parameter space $\Theta$?

   (b) If the null hypothesis is $H_0 : \beta = \beta_0$, what is $\Theta_0$?

   (c) What is $\widehat{\beta}$? Just use your answer from Assignment 2.

   (d) What is $\widehat{\sigma}^2$? Again just use your answer from Assignment 2.

   (e) What is the restricted MLE of $\beta$?

   (f) What is $\widehat{\sigma}_0^2$? Show your work.

   (g) Show $G^2 = n \ln \frac{\sum_{i=1}^n (Y_i - \beta_0 x_i)^2}{\sum_{i=1}^n (Y_i - \widehat{\beta} x_i)^2}$

6. Let $\mathbf{D}_1, \ldots, \mathbf{D}_n$ be a random sample from a multivariate normal population with mean $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$. Write $\mathbf{D}_i = \left( \begin{array}{c} \mathbf{X}_i \\ \hline \mathbf{Y}_i \end{array} \right)$, where $\mathbf{X}_i$ is $q \times 1$, $\mathbf{Y}_i$ is $r \times 1$, and $p = q + r$, we have $cov(\mathbf{D}_i) = \boldsymbol{\Sigma} = \left( \begin{array}{c|c} \boldsymbol{\Sigma}_x & \boldsymbol{\Sigma}_{xy} \\ \hline \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_y \end{array} \right)$, and $\widehat{\boldsymbol{\Sigma}} = \left( \begin{array}{c|c} \widehat{\boldsymbol{\Sigma}}_x & \widehat{\boldsymbol{\Sigma}}_{xy} \\ \hline \widehat{\boldsymbol{\Sigma}}_{yx} & \widehat{\boldsymbol{\Sigma}}_y \end{array} \right)$. We want to test whether the vector of observations $\mathbf{X}_i$ is independent of the vector of observations $\mathbf{Y}_i$. Because zero covariance implies independence for the multivariate normal, the null hypothesis is $H_0 : \boldsymbol{\Sigma}_{xy} = \mathbf{0}$.

   (a) Starting from the formula sheet, write down and simplify the log likelihood evaluated at the unrestricted MLE. Your answer is a formula. It's in the lecture slides if you want to check your answer.

   (b) Using the fact that zero covariance implies independence for the multivariate normal, give the restricted MLE $(\widehat{\boldsymbol{\mu}}_0, \widehat{\boldsymbol{\Sigma}}_0)$. If you think about it you can just write down the answer without any calculation.

(c) Give the log likelihood evaluated at the restricted MLE. It's easy if you use the fact that zero covariance implies independence for the multivariate normal; otherwise, you're essentially re-proving this fact.

(d) Calculate and simplify the large-sample likelihood ratio statistic $G^2$ for testing $H_0 : \mathbf{\Sigma}_{xy} = \mathbf{0}$, which is equivalent to $\mathbf{X}_i$ and $\mathbf{Y}_i$ independent. Start with the likelihood and MLEs on the formula sheet. Your answer is a formula. What are the degrees of freedom?

(e) For example, $\mathbf{X}_i$ could be the vector of three "mental measurements," namely scores on standardized tests of vocabulary and ability to solve puzzles. $\mathbf{Y}_i$ could be a vector of six physical measurements (head circumference etc.). For $n = 74$, I calculated $\ln|\widehat{\mathbf{\Sigma}}| = 40.814949$, $\ln|\widehat{\mathbf{\Sigma}}_x| = 14.913525$ and $\ln|\widehat{\mathbf{\Sigma}}_y| = 26.33133$.

   i. Calculate $G^2$ for these data. Your answer is a number. My answer is also a number: ??.81304.

   ii. What are the degrees of freedom? Your answer is a number.

   iii. The critical value at $\alpha = 0.05$ is not on the formula sheet. It's 28.8693. Do you reject $H_0$? Are the mental and physical characteristics independent?

7. In a study of diet and health, suppose we want to know how much snack food each person eats, and we "measure" it by asking a question on a questionnaire. Surely there will be measurement error, and suppose it is of a simple additive nature. But we are pretty sure people under-report how much snack food they eat, so a model like $W = X + e$ with $E(e) = 0$ is hard to defend. Instead, let

$$W = \nu + X + e,$$

where $E(X) = \mu_x$, $E(e) = 0$, $Var(X) = \sigma_x^2$, $Var(e) = \sigma_e^2$, and $Cov(X, e) = 0$ The unknown constant $\nu$ could be called *measurement bias.* Calculate the reliability of $W$ for this model. Is it the same as the expression for reliability given in the text and lecture, or does $\nu \neq 0$ make a difference?

8. Continuing Exercise 7, suppose that two measurements of $W$ are available.

$$\begin{aligned} W_1 &= \nu_1 + X + e_1 \\ W_2 &= \nu_2 + X + e_2, \end{aligned}$$

where $E(X) = \mu_x$, $Var(X) = \sigma_x^2$, $E(e_1) = E(e_2) = 0$, $Var(e_1) = Var(e_2) = \sigma_e^2$, and $X$, $e_1$ and $e_2$ are all independent. Calculate $Corr(W_1, W_2)$. Does this correlation still equal the reliability even when $\nu_1$ and $\nu_2$ are non-zero and potentially different from one another?

9. Let $X$ be a latent variable, $W = X + e_1$ be the usual measurement of $X$ with error, and $G = X + e_2$ be a measurement of $X$ that is deemed "gold standard," but of course it's not completely free of measurement error. It's better than $W$ in the sense that $0 < Var(e_2) < Var(e_1)$, but that's all you can really say. This is a realistic scenario, because nothing is perfect. Accordingly, let

$$\begin{aligned} W &= X + e_1 \\ G &= X + e_2, \end{aligned}$$

where $E(X) = \mu_x$, $Var(X) = \sigma_x^2$, $E(e_1) = E(e_2) = 0$, $Var(e_1) = \sigma_1^2$, $Var(e_2) = \sigma_2^2$ and $X$, $e_1$ and $e_2$ are all independent of one another.

(a) Make a path diagram of this model.

(b) Prove that the squared correlation between $W$ and $G$ is strictly less than the reliability of $W$. Show your work.

The idea here is that the squared *population* correlation[2] between an ordinary measurement and an imperfect gold standard measurement is strictly less than the actual reliability of the ordinary measurement. If we were to estimate such a squared correlation by the corresponding squared *sample* correlation, we would be estimating a quantity that is not the reliability. On the other hand, we would be estimating a lower bound for the reliability, and this could be reassuring if it were a high number.

10. In this continuation of Exercise 9, show what happens when you calculate the squared *sample* correlation between a usual measurement and an imperfect gold standard and let the sample size $n \to \infty$. It's just what you would think.

11. Suppose we have two equivalent measurements with uncorrelated measurement error:

$$\begin{aligned} W_1 &= X + e_1 \\ W_2 &= X + e_2, \end{aligned}$$

where $E(X) = \mu_x$, $Var(X) = \sigma_x^2$, $E(e_1) = E(e_2) = 0$, $Var(e_1) = Var(e_2) = \sigma_e^2$, and $X$, $e_1$ and $e_2$ are all independent. What if we were to measure the true score $X$ by adding the two imperfect measurements together? Would the result be more reliable?

(a) Let $S = W_1 + W_2$. Calculate the reliability of $S$.

(b) Suppose you take $n$ independent measurements (in psychometric theory, these would be called equivalent test items). What is the reliability of $S = \sum_{i=1}^n W_i$? Show your work.

(c) What is the reliability of $\overline{W}_n = \frac{1}{n}\sum_{i=1}^n W_i$? Show your work.

(d) What happens to the reliability of $S$ and $\overline{W}_n$ as the number of measurements $n \to \infty$?

12. Consider the two equivalent measurements at the start of Question 11. It is easy to imagine omitted variables that would affect both observed scores. For example, if $W_1$ and $W_2$ are two questionnaires about eating habits, some people will probably mis-remember or lie the same way on both questionnaires. Since $e_1$ and $e_2$ represent "everything else," this means that $e_1$ and $e_2$ will have non-zero covariance. Furthermore, this covariance will be positive, since the omitted variables (there could be dozens of them) will tend to affect the two measurements in the same way. Accordingly, in the initial model of Question 11, let $Cov(e_1, e_2) = \kappa > 0$.

(a) Draw a path diagram of the model.

(b) Show that $Corr(W_1, W_2)$ is strictly *greater* than the reliability.

This means that in practice, omitted variables will result in over-estimates of reliability. There are almost always omitted variables.

[2]When we do Greek-letter calculations, we are figuring out what is happening in the population from which a data set might be a random sample.