Statistics 252 "Practice" Midterm #2 Solutions– Winter 2005

1. (a) By definition, the significance level α is the probability of a Type I error; that is, the probability under H_0 that H_0 is rejected. Hence, since $\overline{Y} \sim \mathcal{N}(\mu, 4/n)$,

$$\alpha = P_{H_0}(\text{reject } H_0) = P(\overline{Y} > 3.92/\sqrt{n}|\mu = 0) = P\left(\frac{\overline{Y}-0}{2/\sqrt{n}} > \frac{3.92/\sqrt{n}-0}{2/\sqrt{n}}\right)$$
$$= P(Z > 1.96) = 0.025,$$

where $Z \sim \mathcal{N}(0,1)$. (The last step follows from Table 4.) Hence, we see that the hypothesis test does, in fact, have significance level $\alpha = 0.025$.

1. (b) By definition, the power of an hypothesis test is the probability under H_A that H_0 is rejected. Hence, when $\mu = 0.5$, we find

$$\text{power} = P_{H_A}(\text{reject } H_0) = P(\overline{Y} > 3.92/\sqrt{n}|\mu = 0.5) = P\left(\frac{\overline{Y}-0.5}{2/\sqrt{n}} > \frac{3.92/\sqrt{n}-0.5}{2/\sqrt{n}}\right)$$
$$= P(Z > 1.96 - 0.25\sqrt{n})$$

where $Z \sim \mathcal{N}(0,1)$. If we desire the test to have power 0.9, then using Table 4, we find $P(Z > -1.28) = 0.90$. Thus, we require that n satisfy

$$1.96 - 0.25\sqrt{n} = -1.28 \quad \text{or} \quad n \approx 168.$$

(In fact, we can take $n \geq 168$ to guarantee that the test will have power (at least) 0.9 when $\mu = 0.5$.

2. In this problem, we find that $\alpha = P(\overline{Y} < c|\mu = 0)$ and $\beta = P(\overline{Y} > c|\mu = -1/2)$. Since $\overline{Y} \sim \mathcal{N}(\mu, \sigma^2/n) = \mathcal{N}(\mu, 0.25)$, we conclude that

$$\alpha = P(\overline{Y} < c|\mu = 0) = P\left(\frac{\overline{Y}-0}{\sqrt{0.25}} < \frac{c-0}{\sqrt{0.25}}\right) = P(Z < 2c)$$

and

$$\beta = P(\overline{Y} > c|\mu = -1/2) = P\left(\frac{\overline{Y}+1/2}{\sqrt{0.25}} > \frac{c+1/2}{\sqrt{0.25}}\right) = P(Z > 2c+1)$$

where $Z \sim \mathcal{N}(0,1)$. In order for $\alpha = \beta$, we require that $P(Z < 2c) = P(Z > 2c+1)$. Since the standard normal distribution is symmetric about 0, we see that we must have $-2c = 2c + 1$ or $c = -1/4$. (DRAW A PICTURE TO SEE WHERE THE MINUS SIGN COMES FROM!) Consulting Table 4, we find that with $c = -1/4$, the significance level of the this test is

$$\alpha = P(Z < -1/2) = 0.3085.$$

3. (a) Recall that the generalized likelihood ratio test for the simple null hypothesis $H_0 : \theta = \theta_0$ against the composite alternative hypothesis $H_A : \theta \neq \theta_0$ has rejection region $\{\Lambda \leq c\}$

where Λ is the generalized likelihood ratio

$$\Lambda = \frac{L(\theta_0)}{L(\hat{\theta}_{\text{MLE}})}$$

where $L(\theta)$ is the likelihood function. In this instance,

$$L(\theta) = \theta^n \exp\left(-\theta \sum_{i=1}^{n} y_i\right)$$

so that

$$\Lambda = \frac{\theta_0^n \exp\left(-\theta_0 \sum_{i=1}^{n} y_i\right)}{\hat{\theta}_{\text{MLE}}^n \exp\left(-\hat{\theta}_{\text{MLE}} \sum_{i=1}^{n} y_i\right)} = \left(\frac{\theta_0}{1/\overline{Y}}\right)^n \exp\left(-\theta_0 \sum y_i + 1/\overline{Y} \cdot \sum y_i\right)$$

$$= \left(\theta_0 \overline{Y}\right)^n \exp\left(n - n\theta_0\overline{Y}\right) = e^n\, \theta_0^n\, \overline{Y}^{\,n} \exp\left(-n\theta_0\overline{Y}\right) = e^n\, \theta_0^n \left[\overline{Y} \exp\left(-\theta_0\overline{Y}\right)\right]^n$$

Hence, we see that the rejection region $\{\Lambda \le c\}$ can be expressed as

$$\left\{e^n\, \theta_0^n \left[\overline{Y} \exp\left(-\theta_0\overline{Y}\right)\right]^n \le c\right\} = \left\{\overline{Y} \exp\left(-\theta_0\overline{Y}\right) \le c^{1/n} e^{-1}\, \theta_0^{-1}\right\}$$
$$= \left\{\overline{Y} \exp\left(-\theta_0\overline{Y}\right) \le C\right\}$$

(To be explicit, the *suitable constant* is $C = c^{1/n} e^{-1}\, \theta_0^{-1}$, although this was "not required.")

3. (b) We saw in class that $-2\log\Lambda \sim \chi_1^2$ (approximately). This means that the generalized likelihood ratio test rejection region is $\{\Lambda \le c\} = \{-2\log\Lambda \ge K\}$ where K is (yet another) constant. As we found above,

$$\Lambda = e^n\, \theta_0^n \left[\overline{Y} \exp\left(-\theta_0\overline{Y}\right)\right]^n$$

so that

$$-2\log\Lambda = -2n - 2n\log\theta_0 - 2n\log\overline{Y} + 2n\theta_0\overline{Y}.$$

Hence, to conduct the GLRT, we need to compare the observed value of $-2\log\Lambda$ with the appropriate chi-squared critical value which is $\chi^2_{1,0.10} = 2.70554$. Since

$$-2 \cdot 10 - 2 \cdot 10 \log 1 - 2 \cdot 10 \cdot \log 1.25 + 2 \cdot 10 \cdot 1 \cdot 1.25 \approx 2.76856$$

is the observed value of $-2\log\Lambda$, we reject H_0 at significance level 0.10. (Note, however, that since $\chi^2_{1,0.05} = 3.84146$, we fail to reject H_0 at significance level 0.05. Again, this is for your edification, and was "not required.")

4. (a) Consider an hypothesis test of $H_0 : \theta = \theta_0$ against H_A where H_A could be any one of $H_A : \theta \ne \theta_0$, $H_A : \theta > \theta_0$, or $H_A : \theta < \theta_0$. The significance level α is simply the probability of a Type I error. A Type I error occurs if H_0 is rejected when, in fact, H_0 is true. Thus,

$$\alpha = P(\text{Type I error}) = P_{H_0}(\text{reject } H_0).$$

4. (b) Consider the linear regression model $Y = \beta_0 + \beta_1 x + \varepsilon$ with linear trend $\mathbb{E}(Y) = \beta_0 + \beta_1 x$. The least squares estimators of β_0 and β_1, written $\hat{\beta}_0$ and $\hat{\beta}_1$, yield the least squares line $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x$. Now, $\hat{Y}$ can be used BOTH as an estimator of the parameter $\mathbb{E}(Y)$ when the input is x, AND as a predictor of the random variable Y when the input is x. In the first instance, a confidence interval for the parameter $\mathbb{E}(Y)$ is formed, and in the second instance, a prediction interval for the random variable Y is formed. In addition to the difference in *interpretation* as just given, it is important to note that while both the CI and the PI are centred at $\hat{Y}$, they have different formulæ, and hence have different lengths in general. (See Figure 11.7 on page 566.)

5. The likelihood function (joint density function of the Y_i) is

$$L(\beta) = f(y_1, \ldots, y_n|\beta) = \prod_{i=1}^{n} f(y_i|\beta) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right).$$

In order to maximize the likelihood function, it is equivalent to maximize the log-likelihood function

$$\ell(\beta) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2.$$

Taking derivatives yields

$$\frac{\partial}{\partial\beta}\ell(\beta) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \beta x_i)x_i = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i y_i - \beta\sum_{i=1}^{n} x_i^2\right).$$

Setting

$$\frac{\partial}{\partial\beta}\ell(\beta) = 0$$

and solving gives

$$\sum_{i=1}^{n} x_i y_i - \beta\sum_{i=1}^{n} x_i^2 = 0$$

so that

$$\hat{\beta}_{\text{MLE}} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}.$$

Since this is only a function of one variable, we need to check that this is actually a maximum. Taking second derivatives gives

$$\frac{\partial^2}{\partial\beta^2}\ell(\beta) = -\sum_{i=1}^{n} x_i^2 < 0$$

so that by the second derivative test we do have a maximum. Notice that the maximum likelihood estimator of β is exactly the least squares estimator of β that you found in Exercise 11.6. The computation is, in fact, identical. However, the *interpretation* is quite distinct.