Jack Rodoni
STAT 630
Assignment 11
12/8/21

1.) S: $\mathcal{U}_1, \ldots \mathcal{U}_n$ is a random sample from Unif $(0,\theta)$ where θ is unknown.
[*Recall that $M_n = \max(\mathcal{U}_1, \ldots, \mathcal{U}_n)$ is the MLE for θ. *]

(a) Consider the test for $H_0: \theta = 1$ vs $H_A: \theta = \theta_1$ w/ $\theta_1 > 1$, that rejects H_0 when $M_n > c$ for some $c < 1$. Identify c so that the test has size α. (Recall $M_n \le c \iff$ every $\mathcal{U}_i < c$)

$H_0: \theta = 1$, $H_A: \theta > 1$

- $\alpha = P_{H_0}(M_n > c)$

$\Rightarrow 1-\alpha = P(M_n \le c \mid \theta = 1)$

$= P(\max\{\mathcal{U}_1, \ldots \mathcal{U}_n\} \le c)$

$= (P(\mathcal{U}_i \le c))^n$

$= \left(\int_0^c f_{\mathcal{U}_1}(u \mid \theta = 1)\, du\right)^n$

$= \left[\int_0^c 1\, du\right]^n$

$1-\alpha = c^n \Rightarrow \boxed{c = (1-\alpha)^{1/n}}$

(b) Determine the power of this test as a function of θ,

$\gamma(\theta) = P(\text{reject } H_0, \text{ when } H_1 \text{ is true})$

$\gamma(\theta) = P(M_n > c)$

$= 1 - P_{\theta_1}(M_n < c)$

$= 1 - (P_{\theta_1}(\mathcal{U}_1 < c))^n$

$= 1 - \left[\int_0^c f_{\mathcal{U}_1}(u \mid \theta_1)\, du\right]^n$

$= 1 - \left[\int_0^c \frac{1}{\theta_1}\, du\right]^n$

$= 1 - \left(\frac{c}{\theta_1}\right)^n = 1 - \left(\frac{(1-\alpha)^{1/n}}{\theta_1}\right)^n = \boxed{1 - \left(\frac{1-\alpha}{\theta_1^n}\right) = \gamma(\theta)}$

2.) Exercises 6.3.26, 6.3.27.

Exercise 6.3.26: (One sided Hypothesis for means) $\&$ $(X_1,\dots,X_n)$ is a sample from a $N(\mu,\sigma_0^2)$ distribution, where μ is unknown & σ_0^2 is known. & we want to assess the hypothesis $H_0: \mu \le \mu_0$. Under these circumstances we say that the observed value $\bar{x}$ is surprising if $\bar{x}$ occurs in a region of low probability for every distribution in H_0. Therefore, a sensible P-value for this problem is $\max_{\mu \in H_0} P_\mu(\bar{X} > \bar{x})$ show that this leads to the p-value $1-\Phi((\bar{x}-\mu_0)/(\sigma_0/\sqrt{n}))$

$$P(\bar{X} \ge \bar{x}) = P\left[Z \ge \frac{(\bar{x}-\mu_0)}{s}\right] = P\left[Z \ge \left(\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right)\right] = 1 - P\left[Z \le \left(\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right)\right]$$

$$\boxed{P(\bar{X} \ge x) = 1 - \Phi\left(\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right)}$$

Exercise 6.3.27: Determine the form of the power function associated with the hypothesis assessment procedure of problem 6.3.26, when we declare a test result as being statistically significant whenever the p-value is less than α.

see A.10.6 Testing pg 12

$$\gamma(\mu') = P(\text{Reject } H_0 \text{ when } \mu = \mu')$$
$$= P\left(\bar{X} < \mu_0 - Z_{1-\alpha}\frac{\sigma}{\sqrt{n}} \,\middle|\, \mu = \mu'\right)$$
$$= P_{\mu'}\left(\frac{\bar{X}-\mu'}{\sigma/\sqrt{n}} < \frac{\mu_0-\mu'}{\sigma/\sqrt{n}} - Z_{1-\alpha}\right)$$
$$\gamma(\mu') = \Phi\left(\frac{\mu_0-\mu'}{\sigma/\sqrt{n}} - Z_{1-\alpha}\right)$$

- Also explain why the test described here (namely "reject H_0 if p-value $\le \alpha$") is the same as the test that rejects H_0 when $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge Z_{1-\alpha}$

"explain later"

3.) Chp 8 Exercise 8.2.3: $\$$ an investigator knows that an industrial process yields a response variable that follows a $N(1,2)$ distribution. Some changes have been made to the process and the investigator believes that these have possibly made a change in the mean response (not the variance) increasing its value. The investigator wants the probability of a Type I error occurring to be less than 1%. Determine an appropriate testing procedure for this problem based on a sample size 10.

(a) First identify the Hypothesis & explain why.

- $H_0: \mu \le 1$, $H_a: \mu > 1$

(b) Determine an appropriate testing procedure for this problem based on a sample size 10.

- $1 - P\left(Z_{0.99} \le \frac{\bar{X}-1}{2/\sqrt{10}}\right)$

4.) Chp 6.3.1 & 6.3.2 (Assignment 9). Now compute p-value for each. Do not do a power calculation.

Chp 6 Exercise 6.3.1: Some measurements (in cm) are taken using an instrument. There is error in the measuring process and a measurement is assumed to be distributed $N(\mu, \sigma_0^2)$, where μ is the exact measurment and $\sigma_0^2 = 0.5$. If the $(n=10)$ measurments 4.7, 5.5, 4.4, 3.3, 4.6, 5.3, 5.2, 4.8, 5.7, 5.3 were obtained.

Assess the hypothesis

$H_0: \mu = 5$ $H_a\ \mu \neq 0.5$

Also compute a 95% CI for the unknown μ.

- $\text{p-value} = 2\left[P\left(Z_{.975} \le \frac{4.88-5}{\sqrt{0.5}/\sqrt{10}}\right)\right] \overset{\text{in R}}{=} 2 * \text{pnorm}\left(\frac{4.88-5}{\sqrt{0.5}/\sqrt{10}}\right)$

$\boxed{\text{p-value} = 0.591505 \Rightarrow \text{fail to reject } H_0}$

- CI: $\bar{X} \pm Z_{.975} \frac{\sigma}{\sqrt{n}} = 4.88 \pm 1.959964\left(\frac{\sqrt{0.5}}{\sqrt{10}}\right)$

$\boxed{CI = (4.441739, 5.318261)}$

4.) Contd.

t-test

Exercise 6.3.2: β in exercise 6.3.1, we drop the assumption that $\sigma_0^2 = 0.5$

Then assess the hypothesis

$H_0: \mu = 5$, $H_1: \mu \neq 5$

and compute a 0.95 confidence interval for μ.

- p-value $= 2\left[P\left[t_{0.975} \le \frac{4.88-5}{s/\sqrt{10}}\right]\right] \overset{\text{in R}}{=} 2 * pt\left(\frac{4.88-5}{s/\sqrt{10}}, df=9\right)$

$\boxed{\text{p-value} = 0.5986097 \Rightarrow \text{fail to reject } H_0}$

- CI: $\bar{x} \pm t_{0.975}\frac{s}{\sqrt{n}} = (4.382325, 5.377675)$

$\boxed{\text{CI}: (4.382325, 5.377675)}$

5.) Chap 6 Exercise 6.3.8 (Assignment 9): Now compute p-values using both the Wald & score statistics.

A polling firm conducts a poll to determine what proportion θ of voters in a given population will vote in an upcoming election. A random sample of $n=250$ was taken from the population and the proportion answering yes was 0.62. Assess the hypothesis

$H_0: \theta = 0.65$, $H_1: \theta < 0.65$

and construct an approximate 0.90 confidence interval for θ.

- Wald: $z = \hat{\theta} - \theta_0 \Big/ \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$, p-value $=$ pnorm(z) $= 0.1642238$

* ~~get CI from HW 9~~

- Score: $S = \frac{n(\hat{\theta}-\theta_0)^2}{\theta_0(1-\theta_0)}$, p-value $= 1 -$ pnorm(sqrt(S)) $= 0.1599921$

6.) $X \sim \text{binomial}(100, \theta)$. Consider the test that rejects $H_0: \theta = 0.5$ in favor of $H_a: \theta \neq 0.5$ when $|X - 50| > 10$.

Use the normal approximation to answer the following:

(a) What is α?

$$\boxed{\alpha = \text{pnorm}(40, 50, 5) + (1 - \text{pnorm}(60, 50, 5)) = 0.04550026}$$

where $50 = np$; $5 = \sqrt{np(1-p)}$

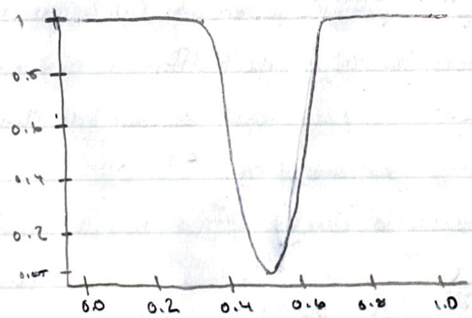
(b) Derive the (approximate power function & graph it)

$$\gamma(\hat{\theta}) = \Phi\left[\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} - Z_{1-\alpha/2}\right] + \left(1 - \Phi\left[\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} + Z_{1-\alpha/2}\right]\right)$$

$$= \Phi\left[\frac{\sqrt{n}(\hat{\theta} - \theta_0)}{\sqrt{\hat{\theta}(1-\hat{\theta})}} - Z_{1-\alpha/2}\right] + \left(1 - \Phi\left[\frac{\sqrt{n}(\hat{\theta} - \theta_0)}{\sqrt{\hat{\theta}(1-\hat{\theta})}} + Z_{1-\alpha/2}\right]\right)$$

$$= \Phi\left[\frac{\sqrt{100}(\hat{\theta} - 0.5)}{\sqrt{\hat{\theta}(1-\hat{\theta})}} - Z_{1-\alpha/2}\right] + \left(1 - \Phi\left[\frac{\sqrt{100}(\hat{\theta} - 0.5)}{\sqrt{\hat{\theta}(1-\hat{\theta})}} + Z_{1-\alpha/2}\right]\right)$$

Plot-



7.) Chp 8 Exercise 8.2.4: If you have a sample of 20 from a $N(\mu, 1)$ distribution

You form a 0.975 - CI for μ & use it to test $H_0: \mu = 0$ by rejecting H_0 whenever

0 is not in the confidence interval.

(a) What is the size of this test?

$\boxed{\alpha = 1 - 0.975 = 0.025}$

(b) Determine the power function of this test.

$$\gamma(\mu) = \Phi\left[\frac{\mu' - \mu_0}{\sigma/\sqrt{n}} - Z_{1-\alpha/2}\right] + \left(1 - \Phi\left[\frac{\mu' - \mu_0}{\sigma/\sqrt{n}} + Z_{1-\alpha/2}\right]\right)$$

$$\boxed{\gamma(\mu) = \Phi\left[\frac{\mu' - \mu}{1/\sqrt{20}} - Z_{0.9875}\right] + \left(1 - \Phi\left[\frac{\mu' - \mu}{1/\sqrt{20}} + Z_{0.9875}\right]\right)}$$

8.) Chp 8 Exercise 8.2.6: No computation is necessary. Just think about what "having not ruled the null hypothesis out" signifies here.

If: you are testing a null hypothesis $H_0: \theta = 0$ where $\theta \in R^1$. You use a size $\alpha = 0.05$ testing procedure and accept H_0. You feel you have a fairly large sample, but when you compute the power at ± 0.2, you obtain a value of 0.10 where 0.2 represents the smallest difference that is of practical importance. Do you believe it makes sense to conclude that the null hypothesis is true? Justify your conclusion.

• No, it does not make sense to conclude that the null hypothesis is true b/c the power of our test is so low. it is unlikely we would pick up a true deviation from our hypothesized value.

10 (37-3

a.) Recall Exercise 6.2.19 (Assignment 8) for which you found the likelihood and score functions & the MLE. Now, determine the Wald & score tests for a two-sided size α test of $H_0: \theta = 0.5$, $H_a: \theta \neq 0.5$. (That is the two alleles are equally likely (A or a) and are independent.

NOTE: in HW 8 we found:

$$S(\theta|s) = \frac{\partial}{\partial\theta}\left[\ell(\theta|s)\right] = \frac{2x_1 + x_2}{\theta} - \frac{x_2 + 2x_3}{(1-\theta)}$$

$$\hat{\theta} = \frac{2x_1 + x_2}{2n}$$

$$\left[\text{• Recall Fishers information: } I(\theta) = E_\theta\left[\left(\frac{\partial}{\partial\theta}\left[\ln(f_\theta(x))\right]\right)^2\right] = -E_\theta\left[\left(\frac{\partial^2}{\partial\theta^2}(\ln(f_\theta(x)))\right)\right]\right]$$

$$\text{• } \frac{\partial^2}{\partial\theta^2}\ln(f_\theta(x)) = \frac{\partial}{\partial\theta}\left[S(\theta|s)\right] = -\frac{(2x_1 + x_2)}{\theta^2} - \frac{x_2 + 2x_3}{(1-\theta)^2}$$

$$\text{• } I(\theta) = -E\left[-\left(\frac{2x_1 + x_2}{\theta^2}\right) - \left(\frac{x_2 + 2x_3}{(1-\theta)^2}\right)\right] = E\left[\frac{2x_1 + x_2}{\theta^2} + \frac{x_2 + 2x_3}{(1-\theta)^2}\right]$$

$$= \frac{1}{\theta^2}E\left[2x_1 + x_2\right] + \frac{1}{(1-\theta)^2}E\left[x_2 + 2x_3\right]$$

$$\left[\text{• } E[x_1] = \theta^2 n, \quad E[x_2] = 2(\theta)(1-\theta)n, \quad E[x_3] = (1-\theta)^2 n\right]$$

$$= \frac{1}{\theta^2}\left[2(\theta^2 n) + 2(1-\theta)(\theta)n\right] + \frac{1}{(1-\theta)^2}\left[2(1-\theta)(\theta)n + 2(1-\theta)^2 n\right]$$

$$= 2n + \frac{2(1-\theta)n}{\theta} + \frac{2\theta n}{1-\theta} + 2n = 4n + \frac{2n(1-\theta)^2 + 2n\theta^2}{\theta(1-\theta)}$$

$$= \frac{4n\theta(1-\theta) + 2n(1-\theta)^2 + 2n\theta^2}{\theta(1-\theta)} = \frac{4n\theta - 4n\theta^2 + 2n(1-2\theta+\theta^2) + 2n\theta^2}{\theta(1-\theta)}$$

$$= \frac{\cancel{4n\theta} - \cancel{4n\theta^2} + 2n - \cancel{4n\theta} + \cancel{2n\theta^2} + \cancel{2n\theta^2}}{\theta(1-\theta)} = \boxed{\frac{2n}{\theta(1-\theta)} = I(\theta)} \;*$$

$$\text{• } I(\hat{\theta}) = \frac{2n}{\hat{\theta}(1-\hat{\theta})} = \frac{2n}{\left(\frac{2x_1+x_2}{2n}\right)\left(1 - \frac{2x_1+x_2}{2n}\right)} = \boxed{\frac{8n^3}{(2x_1+x_2)(2n-2x_1-x_2)}} \;*$$

Wald:

$$\boxed{z = \frac{\left(\frac{2x_1+x_2}{2n}\right) - 0.5}{\sqrt{\frac{(2a+b)(2n-2a-b)}{8n^3}}}}$$

Score:

$$z = \frac{\left(\frac{2x_1 + x_2}{\theta_0}\right) - \left(\frac{x_2 + 2x_3}{(1-\theta_0)}\right)}{\sqrt{\frac{(1-\theta_0)\theta_0}{2n}}} \overset{\theta_0 = 0.5}{=} \frac{2(2x_1 + x_2 - x_2 + 2x_3)}{\sqrt{\frac{(0.5)^2}{2n}}}$$

$$\boxed{z = 8(x_1 + x_3)\sqrt{2n}}$$

(See H.O.6 (testing) pg 37-39)

10.) Ch. 8 Exercise 8.2.20. Express the test in terms of a sufficient statistic (and its sampling distribution.) (While it gives a way to express the power function as an integral, you can answer both parts (a) & (b) without the hint by using the Neyman-Pearson Lemma and what we have said about what makes a UMP test.)

- S that $(X_1, \ldots X_n)$ is a sample from a poisson (λ) distribution where $\lambda > 0$ is unknown.

(a) Determine the UMP size α test for $H_0: \lambda = \lambda_0$ $H_a: \lambda = \lambda_1$ where $\lambda_0 < \lambda_1$.

$H_0: \lambda = \lambda_0$ vs $H_a: \lambda > \lambda_0$

(b) Is this test function UMP size α for $H_0: \lambda \leq \lambda_0$ $H_a: \lambda > \lambda_0$

Yes, they are going to yield the same results given the same α.

(c) Derive the generalized likelihood ratio for testing $H_0: \lambda = \lambda_0$ $H_a: \lambda \neq \lambda_0$ and describe how to use it to conduct the test.

- $LR = \dfrac{\prod f_{\lambda_1}(x_i)}{\prod f_{\lambda_0}(x_0)} = \dfrac{\prod_{i=1}^{n} \frac{\lambda_1^{x_i} e^{-\lambda_1}}{x_i}}{\prod_{i=1}^{n} \frac{\lambda_0^{x_i} e^{-\lambda_0}}{x_i}} = \dfrac{\prod \lambda_1^{x_i} e^{-\lambda_1}}{\prod \lambda_0^{x_i} e^{-\lambda_0}} = e^{-n(\lambda_1 - \lambda_0)} \left(\dfrac{\lambda_1}{\lambda_0}\right)^{n\lambda_1}$

- $2 \ln(LR) = -2n(\lambda_1 - \lambda_0) + 2n\lambda_1 \ln\left(\frac{\lambda_1}{\lambda_0}\right)$

rejection region: $2 \log(LR) \geq \chi^2_{1-\alpha}(1)$

- To use we would sub in the mle of λ under the alternative and make our decision based on the rejection region.

(d) Use the data of problem 3, Assignment 9 to test the Hypothesis:

$\lambda_0 = 2$ vs $\lambda_0 \neq 2$ w/ $\alpha = 0.01$

- $\hat{\lambda} = 1.68$, $n = 150$

$2 \log(LR) = -2(150)(1.68 - 2) + 2(150)(1.68) \ln\left(\frac{1.68}{2}\right)$

$= 8.125893$

$\chi^2_{0.99}$ = qchisq(0.99, 1) = 6.634

$8.125893 \geq 6.634 \Rightarrow$ reject H_0.