Worksheet 11 - Composite Hypotheses and Extensions of the Maximum Likelihood Ratio

Shawn Kim

***Directions**: Please upload a PDF to Gradescope that includes both your written responses and corresponding R code inputs/outputs (if requested) for each problem.*

Problem 1. Younger Americans are better than their elders at separating factual from opinion statements in the news, according to a new analysis from Pew Research Center. Included in their study were 755 adults between the ages of 18 and 29.

Problem 1 Part a) For adults between the ages of 18 and 29, 34% were able to correctly classify 5 factual and 5 opinion statements. Create and interpret a 95% confidence interval for the proportion of Americans from the age group who would make the correct classification.

NOTE: To do this task, enter your code in the R chunk below by filling in the blanks denoted with `FILL IN` *and uncommenting all non-NOTE lines.*

```
# NOTE: Define the given information
p_hat <- 0.34
n <- 755

# NOTE: Compute the z_star value for 95% confidence interval
z_star <- qnorm(0.975)

# NOTE: Now we can develop the confidence interval. You will need to FILL IN
# the appropriate center and standard deviation

CI <- p_hat + c(-z_star, z_star) * sqrt(p_hat * (1 - p_hat)/n)
round(CI, 4)
```

```
## [1] 0.3062 0.3738
```

This means that we are 95% confident that the average percentage of people that successfully distinguish facts from opinions falls within the interval state above, (given repeated samples of the same size n = 755, 95% of them would contain the true mean)

Problem 1 Part b) An educational program aimed at youth is designed to help improve the ability to classify a fact from an opinion. With a composite hypothesis

$$H_0 : p \leq p_0 \quad \text{versus} \quad H_1 : p > p_0$$

and for $p_0 = 0.34$, determine the value (rounded to four decimal places) of the power function $\pi(p)$ for $p = 0.34, 0.35, 0.36, 0.37.0.38$ and 0.39 with the choice of $\alpha = 0.02$ and a sample of size 755.

Recall: In composite hypothesis testing, we do not know the value p for the alternative distribution, so we analyze the power of the test under possible values p of the alternative. We do this by first determining the critical value $k(p_0)$, then calculating the power (the probability of observing $k(p_0)$ under the alternative distribution with p.)

Hint: Since we have a "greater than" alternative hypothesis, then the power will be the right-end probability. We can calculate this using the cumulative distribution function of the standard normal:

$$\text{power} = P\left(Z > \frac{k(p_0) - p}{\sqrt{p(1-p)/n}}\right) = 1 - P\left(Z < \frac{k(p_0) - p}{\sqrt{p(1-p)/n}}\right) = 1 - \Phi\left(\frac{k(p_0) - p}{\sqrt{p(1-p)/n}}\right)$$

NOTE: To do this task, enter your code in the R chunk below by filling in the blanks denoted with `FILL IN` *and uncommenting all non-NOTE lines.*

```
# NOTE: Define the given information

alpha = 0.02  # significance level

p0 = 0.34  # proportion under null distribution

n = 755  # sample size

# NOTE: Determine critical value (remember, we reject H0 if p_hat > p_crit)

z_crit = qnorm(1 - alpha)  # z_crit at alpha = 0.02

p_crit = p0 + z_crit * sqrt(p0 * (1 - p0)/n)  # critical value under null distribution

# NOTE: Measure power for p_alt = 0.34, 0.35, ..., 0.39 define the possible
# proportions under alt. distribution

p_alt = seq(from = 0.34, to = 0.39, by = 0.01)

# NOTE: calculate power = P(X > p_crit) under alt. distribution you should get
# 0.0200, 0.0716, ..., 0.7945

pwr = 1 - pnorm(p_crit, p_alt, sqrt(p_alt * (1 - p_alt)/n))

# NOTE: Display data.frame of prop under alt. distribution vs power. Power
# should be rounded to 4 decimal places

data.frame(p = p_alt, Power = pwr)
```

```
##      p      Power
## 1 0.34 0.02000000
## 2 0.35 0.07164776
## 3 0.36 0.18890327
## 4 0.37 0.37915366
## 5 0.38 0.60257700
## 6 0.39 0.79449297
```

Problem 1 Part c) What qualitative change would you see in the power curve change if α is reduced to 0.01? Support your answer with a clear explanation.

Hint: If α is reduced, then Type I Error is also reduced. Does this increase/decrease the critical value? How does this new critical value effect the power?

NOTE: Enter your response as text below.

The power curve would be less steep (slope increasing at a slower rate) if the alpha was reduced to 0.01. This makes sense because lower alpha means higher critical value and thus lower chance of type 1 error, signified by the area between critical value and where PDF of null hypothesis approaches 0 on the right side in this case where palt > p0. Thus a critical value further to the right of the null hypothesis is also relatively further right along the curve of the alternative hypothesis, reducing the power of the alternative hypothesis since power is area between critical value and where the pdf of the alternative hypothesis approaches 0 on the right side. You can also use intuition; to have higher confidnce level means that we need tighter bounds so it makes sense that the power of the alternative hypothesis decreases if we want to be more sure about something

Problem 2. Snell's Law tell us how light bends at an interface - the angle of incidence versus the angle of refraction - based on the ratio of the velocities of light in the two isotropic media. If the angle of incidence of a laser beam in a vacuum is θ_1 radians and the angle of refraction in an unknown medium is θ_2 radians, then n is called the **index of refraction**, and is given by:

$$n = \frac{\sin(\theta_1)}{\sin(\theta_2)}.$$

You make 16 repeated independent measurements (in radians) of the angle of incidence in a vacuum, $\theta_{1,1}, \theta_{1,2}, \cdots, \theta_{1,16}$, and 16 repeated independent measurements (in radians) of the angle of refraction in the second medium, $\theta_{2,1}, \theta_{2,2}, \cdots, \theta_{2,16}$. These measurements are described below:

- In vacuum: sample mean $\hat{\mu_{\theta_1}} = \overline{\theta_1} = 0.786$ radians and standard deviation $\sigma_{\theta_1} = 0.03$ radians.
- In 2nd medium: sample mean $\hat{\mu_{\theta_2}} = \overline{\theta_2} = 0.326$ radians and standard deviation $\sigma_{\theta_2} = 0.06$ radians.

Problem 2 Part a) Snell's Law gives an estimate $\hat{n} = n(\overline{\theta}_1, \overline{\theta}_2)$ based on the values of $\overline{\theta}_1$ and $\overline{\theta}_2$.
Apply the Delta Method to obtain the numerical estimate of the mean $E[\hat{n}]$ and standard deviation $\text{SD}(\hat{n})$.

Hint: Recall that by the Delta Method, the estimator $\hat{n} = n(\overline{\theta}_1, \overline{\theta}_2)$ of the index of refraction n is approximately Normal with mean

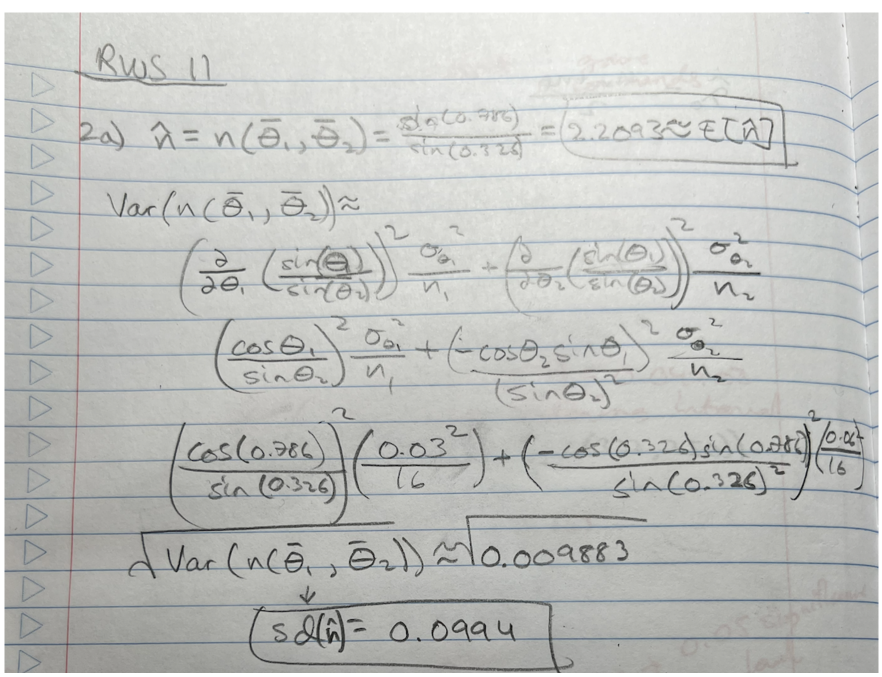
$$E[\hat{n}] = E[n(\overline{\theta}_1, \overline{\theta}_2)] \approx n(\mu_{\theta_1}, \mu_{\theta_2})$$

and variance

$$\text{Var}(\hat{n}) = \text{Var}(n(\overline{\theta}_1, \overline{\theta}_2)) \approx \left[\frac{\partial}{\partial \theta_1} n(\mu_{\theta_1}, \mu_{\theta_2})\right]^2 \cdot \frac{\sigma_{\theta_1}^2}{n_1} + \left[\frac{\partial}{\partial \theta_2} n(\mu_{\theta_1}, \mu_{\theta_2})\right]^2 \cdot \frac{\sigma_{\theta_2}^2}{n_2}$$

if random variables θ_1 and θ_2 are measured n_1 and n_2 independent times, respectively.If calculated correctly, you should conclude $E[\hat{n}] \approx 2.2093$ and $\text{SD}(\hat{n}) \approx 0.0994$.

NOTE: To complete this task, upload an image of your work by replacing "upload_image.jpg" with your appropriately titled .jpg file in the R chunk below.



Problem 2 Part b) You suspect that the substance is cubic zirconia ($n_z = 2.165$) and not diamond, the claimed material, ($n_d = 2.418$). Consequently, you construct the hypothesis

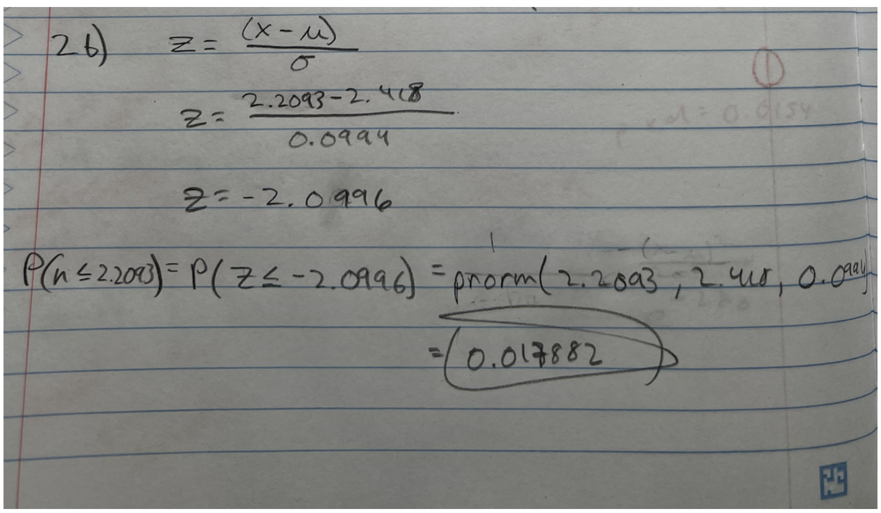
$$H_0 : n = n_d \ (= 2.418) \quad \text{versus} \quad H_1 : n = n_z \ (= 2.165).$$

Devise a z-test for the hypothesis and compute the p-value for this test.

Hint: Note that the sampling distribution of $\hat{n}$ for the null hypothesis has mean 2.418 and (approximate) standard deviation 0.0994, while the sampling distribution of $\hat{n}$ for the alternative hypothesis has mean 2.165 and (approximate) standard deviation 0.0994. (See visualization below.)

Hint: To devise a z-test, you will need to standardize the data under the null hypothesis distribution. The resulting z-score is the z-test statistic. The p-value is the probability of observing the data under the appropriate tail of the null distribution. If done correctly, $P_{n_0}(n \leq 2.2093) = P(Z \leq -2.0996) = 0.0179$.

NOTE: To complete this task, upload an image of your work by replacing "upload_image.jpg" with your appropriately titled .jpg file in the R chunk below.



Problem 2 Part c) Interpret the meaning of the p-value for this test, in context of the application at hand.

NOTE: To complete this task, complete the sentence below with the appropriate number/word in the two blanks. Hint: The choices for the first blank will be diamond or cubic zirconia, and for the third blank the choices will be larger or smaller.

"The p-value means that if the unknown substance really was diamond, then only 1.79 % of the time would we expect to get data with a mean of 2.2093, or smaller, when we run this test."