Valeria Duran

Math 6358

Midterm 2

1a. $\hat{p} = {}^{55}/_{100} = 0.55$

The 95% confidence interval for p is:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Where $z_{0.975} = 1.96$

$$0.55 \pm 1.96 \sqrt{\frac{0.55(1-0.55)}{100}}$$

$$0.55 \pm 0.09750877$$

$$(0.453, 0.648)$$

1b.$\hat{p} - \hat{q} = \hat{p} - (1-\hat{p}) = 2\hat{p} - 1$
$\hat{p} = 0.10$

1c.

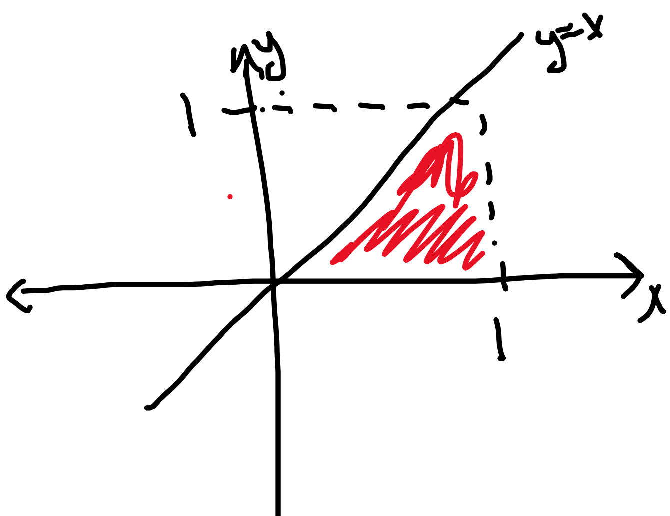
$$0.10 \pm 1.96 \sqrt{\frac{0.10(1-0.10)}{100}}$$

$$SE(\hat{p}) = 0.03$$

$$0.10 \pm 0.0588$$

$$(0.0412, 0.1588)$$

2a.



Limit of integration is from 0 to 1.

2b. $\int_0^1 \int_0^x f(x,y)\, dy\, dx = 1$

$\therefore \int_0^1 \int_0^x k(x-y)\, dy\, dx = 1 \Rightarrow \int_0^1 \left[k\left(xy - \frac{y^2}{2}\right)\right]_0^x dx = 1$

$\Rightarrow \int_0^1 k\left(x^2 - \frac{x^2}{2}\right) dx = 1 \Rightarrow \int_0^1 \frac{kx^2}{2}\, dx = 1$

$\Rightarrow k \int_0^1 \frac{x^2}{2}\, dx = 1 \Rightarrow k\left[\frac{x^3}{6}\right]_0^1 \Rightarrow k\left[\frac{1}{6} - 0\right] = 1$

$\Rightarrow \frac{k}{6} = 1 \Rightarrow k = 6$

2c. $f(x,y) = 6(x-y), \quad 0 \le y \le x \le 1$
$= 0, \quad \text{otherwise}$

$f_X(x) = \int_0^x 6(x-y)\, dy$

$= 6\left(xy - \frac{y^2}{2}\right)_0^x \Rightarrow 6\left(x^2 - \frac{x^2}{2}\right)$

$f_X(x) = 6\,\frac{x^2}{2}$

$\Rightarrow$ $f_X(x) = 3x^2, \quad 0 \le x \le 1$
0, otherwise

$$f_y(y) = \int_y^1 6(x-y)\,dx = 6\left(\frac{x^2}{2} - yx\right)\Big|_y^1$$

$$\Rightarrow 6\left[\frac{1}{2} - y - \frac{y^2}{2} - y^2\right] \Rightarrow 3\left[y^2 - 2y + 1\right]$$

$$\therefore \boxed{\begin{array}{l} f_y(y) = 3(y-1)^2, \quad 0 \le y \le 1 \\ \qquad\qquad 0, \quad \text{otherwise} \end{array}}$$

2d. $\left(3(y-1)^2\right)\left(3x^2\right)$

$$= 9x^2(y-1)^2$$

$$\boxed{\text{Since } f_y(y) \cdot f_x(x) \neq f(x,y) \rightarrow \text{not independent}}$$

3a. $\bar{x} = 0.214$
$s = 0.036$

$$\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$$

> qt(.95, 15)
[1] 1.75305

$$0.214 \pm 1.75305 \frac{0.036}{\sqrt{16}}$$

90% Confidence Interval:

$$(0.1982225, 0.2297774)$$

We are 90% that the true mean lies between 0.198 and 0.229.

3b. Assume it's a random sample from a normal population since the sample size is small.

4a.

Since independent $\rightarrow$

$$f(x_n; \lambda, \theta) = \lambda^n e^{-\lambda \sum_{j=1}^{n} (x_j - \theta)}$$

$$\therefore f(x_n; \lambda, \theta) = \begin{cases} \lambda^n e^{-\lambda \sum_{i=1}^{n} (x_j - \theta)}, & \min(x_j) \geq \theta \\ 0, & \text{otherwise} \end{cases}$$

Maximum likelihood estimator for θ:

$$\frac{d}{d\theta}\left[\ln[f(x_n; \lambda, \theta)]\right] = \frac{d}{d\theta}\left[n \cdot \ln(\lambda) - \lambda \cdot n \cdot \bar{x} + \lambda \cdot n \cdot \theta\right]$$

$$= \lambda n$$

max value must occur at $\boxed{\tilde{\theta} = \min(x_i)}$ since n and λ are positive

Maximum likelihood estimator for λ:

$$\frac{d}{d\lambda}\left[\ln[f(x_n; \lambda, \theta)]\right] = \frac{d}{d\lambda}\left[n \cdot \ln(\lambda) - \lambda \cdot n \cdot \bar{x} + \lambda \cdot n \cdot \theta\right]$$

$$= \frac{n}{\lambda} - n \cdot \bar{x} + n \cdot \theta$$

$$= \frac{n}{\lambda} - n \cdot \bar{x} + n \cdot \min(x_i) = 0$$

$$\frac{1}{\lambda} = \bar{x} - \min(x_i) \Rightarrow \lambda = \frac{1}{\bar{x} - \min(x_i)}$$

$$\boxed{\hat{\lambda} = \frac{1}{\bar{x} - \min(x_i)}}$$

4b. since smallest value is 0.64 $\Rightarrow$

$$\boxed{\hat{\theta} = \min(x_i) = 0.64}$$

$$\hat{\lambda} \Rightarrow \quad \bar{x} = \frac{55.8}{10} = 5.58$$

$$\boxed{\hat{\lambda} = \frac{1}{5.58 - 0.64} \approx 0.2024}$$

5a.

$$E(\hat{\theta}) = E(\bar{X} + 2.33S)$$
$$E(\hat{\theta}) = E(\bar{X}) + 2.33E(S)$$
$$E(\hat{\theta}) = \mu + 2.33\sigma$$

$$\hat{\theta} = \bar{X} + 2.33S$$

5b.

$$V(\hat{\theta}) = V(\bar{X} + 2.33S)$$
$$V(\hat{\theta}) = V(\bar{X}) + 2.33V(S)$$
$$V(\hat{\theta}) = \frac{\sigma^2}{n} + 2.33^2 * \frac{\sigma^2}{2n}$$

$$V(\hat{\theta}) \approx \frac{3.7145 * \sigma^2}{n}$$

$$\sigma_{\hat{\theta}} = \sqrt{\frac{3.7145 * \sigma^2}{n}}$$

$$\sigma_{\hat{\theta}} = \frac{1.9273 * \sigma}{\sqrt{n}}$$

$$\widehat{\sigma_{\hat{\theta}}} \approx \frac{1.9273 * \sigma}{\sqrt{n}}$$

The estimated standard error is the standard deviation.

5c.

$$H_0: \hat{\theta} = 6.75$$
$$H_a: \hat{\theta} < 6.75$$

5d.

$$z = \frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}}$$

$$z = \frac{\bar{X} + 2.33S - 6.75}{\frac{1.9273 * \sigma}{\sqrt{n}}}$$

$$z = \frac{\bar{X} + 2.33\sigma - 6.75\sqrt{n}}{1.9273 * \sigma}$$

$$z = \frac{6.33 + 2.33 * 0.16 - 6.75\sqrt{64}}{1.9273 * 0.16}$$

```
> ((6.33+2.33*0.16-6.75)*sqrt(64))/(1.9273*0.16)
[1] -1.224511
```

$$z = -1.224511$$

Z critical:
[1]-2.326348

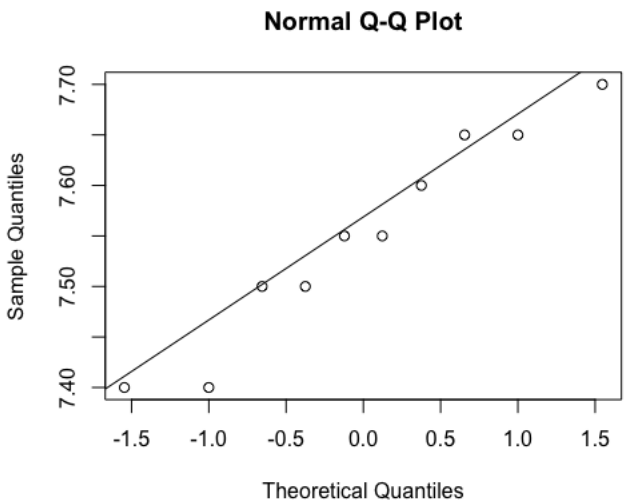
$$-1.224511 > -2.326348$$

We fail to reject the null hypothesis. There is enough evidence to conclude that at most 99% of all possible samples would have a pH of less than 6.75.

6a. Assumptions for a t-test:

1. An SRS of size n form the population
2. The population standard deviation is unknown
3. Assume a normal population since the sample size is small

```
> gum = c(7.65,7.6,7.65,7.7,7.55,7.55,7.4,7.4,7.5,7.5)
> qqnorm(gum)
> qqline(gum)
```




```
> shapiro.test(gum)

        Shapiro-Wilk normality test

data:  gum
W = 0.93973, p-value = 0.55
```

Since the p-value in the Shapiro-Wilk normality test is greater than 0.05, it shows that the sample does follow a normal distribution, as can also be seen by the QQ plot.

6b.

$$H_0: \mu = 7.5$$

$$H_a: \mu \neq 7.5$$

Two-tailed test.

Assume 95% confidence interval will be used

```
> t.test(gum, mu=7.5, conf.level = .95)

        One Sample t-test

data:  gum
t = 1.539, df = 9, p-value = 0.1582
alternative hypothesis: true mean is not equal to 7.5
95 percent confidence interval:
 7.476504 7.623496
sample estimates:
mean of x
     7.55


> qt(.975,9)
[1] 2.262157
```

Since 1.539 < 2.262157, we fail to reject the null hypothesis. Apart from this, the p-value of 0.1582 is greater than the significance level of 0.05, also showing that we fail to reject the null hypothesis. Therefore, we can conclude that the average thickness of gum is equal to 7.5.

We are 95% confident that the true mean of the thickness of gum lies between 7.48 and 7.62.

6c. A Type I error would be rejecting the null hypothesis when it is true. Therefore, a Type I error in this case would be concluding that the average thickness of gum is not 7.5 hundredths of an inch, when in fact it truly is 7.5 hundredths of an inch.

A Type II error would be failing to reject the null hypothesis when it is false. In this case, a Type II error would be concluding that the average thickness of gum is 7.5 hundredths of an inch, when in fact it truly is not 7.5 hundredths of an inch.