Partial exam 21/01/2019

Subject .... Date ....

Assume that $x_1, x_2 — x_n$ are independent realizations of random variable X with density function

$$f_\theta(x) = \frac{\theta x^{\theta-1}}{5^\theta} \quad \text{for } 0 \leq x \leq 5 \qquad \theta > 0$$

(a) calculate the cumulative distribution function of x.

$$F_\theta(x) = \int_0^x \frac{\theta x^{\theta-1}}{5^\theta} dx = \frac{1}{5^\theta} \int_0^x \theta x^{\theta-1} dx = \left[\frac{1}{5^\theta} x^\theta\right]_0^x$$

$$= \frac{1}{5^\theta} x^\theta \implies F_x = \begin{cases} 0 & x < 0 \\ \frac{x^\theta}{5^\theta} & 0 < x < 5 \\ 1 & x > 5 \end{cases}$$

(b) calculate the expected value $\mu = E(X)$ of X.

$$E(X) = \int_0^5 x \frac{\theta x^{\theta-1}}{5^\theta} dx = \frac{1}{5^\theta}\left[\frac{\theta}{\theta+1} x^{\theta+1}\right]_0^5 = \frac{\theta}{\theta+1} \frac{1}{5^\theta}\left[5^{\theta+1} - 0^{\theta+1}\right]$$

$$= \boxed{\frac{5\theta}{\theta+1}}$$

(C) there exists a value of Parameter θ such that the distribution of X is uniform over the interval [0, 5] ?

$$\text{if } \theta = 1 \longrightarrow f_\theta(x) = \frac{1(x)^{1-1}}{5^1} = \frac{1}{5}$$

(d) obtain the likehood function $L(\theta)$.

$$L(\theta) = \prod_{i=1}^n P_{x_i}(x_i, \theta) = \prod_{i=1}^n \frac{\theta x_i^{\theta-1}}{5^\theta}$$

Subject .... Date ....

(e) obtain the log-likehood function $\ell(\theta)$

$$\ell(\theta) = \log(L(\theta)) = \log\left(\prod_{i=1}^{n} \frac{\theta\, x_i^{\theta-1}}{5^{\theta}}\right) = \sum_{i=1}^{n}\left(\log(\theta) + (\theta-1)\log(x_i) - \theta\log 5\right)$$

$$= n\log\theta - n\theta\log 5 + \sum_{i=1}^{n}(\theta-1)\log(x_i)$$

(f) obtain the score function $\ell_*(\theta)$

$$\ell_*(\theta) = \frac{d}{d\theta}(\ell(\theta)) \Rightarrow \ell_*(\theta) = n\frac{1}{\theta} - n\log 5 + \sum_{i=1}^{n}\log(x_i)$$

(g) compute the maximum likehood estimate $\hat{\theta}$ for θ.

$$\ell_*(\theta) = 0 \Rightarrow n\frac{1}{\theta} - n\log 5 + \sum_{i=1}^{n}\log(x_i) = 0 \Rightarrow n\log 5 - \sum_{i=1}^{n}\log(x_i) = n\frac{1}{\theta}$$

$$\boxed{\hat{\theta} = \frac{n}{n\log 5 - \sum_{i=1}^{n}\log(x_i)}}$$

(h) compute $\hat{\theta}$ with the data (see the R-Script below)

1	2	3	4	5	6	7	8	9	10
2.49	4.26	1.72	3.21	2.32	2.38	2.97	4.32	2.40	2.52

```
> length(x)
[1] 10
> sum(log(x))
[1] 4.3995
```

$$(g) \Rightarrow \theta_{ML} = \frac{n}{n\log 5 - \sum_{i=1}^{n}\log(x_i)} = \frac{10}{10\times 0.698 - 4.3995} = \boxed{3.86}$$

$\log 5 = 0.698$

(i) obtain an approximation for the distribution of $\hat{\theta}$.

$$J(\theta) = -\frac{d^2\ell(\theta)}{d\theta^2} = \frac{d}{d\theta}\ell_*(\theta) = +n\frac{1}{\theta^2} \qquad J(\hat{\theta})^{-1} = \frac{\theta^2}{n} \qquad \hat{\theta} \sim N(\theta, J(\theta)^{-1})$$

$$\hat{\theta} \sim N\left(\theta, \frac{\theta^2}{n}\right)$$

Subject.................. Date ....

(J) compute the maximum likehood estimate $\hat{\mu}$ of μ.

$$\hat{\mu} = \frac{5\hat{\theta}}{\hat{\theta}+1} = \frac{5 \times 3.86}{3.86+1} = 3.9711$$

(K) obtain an approximate 95% confidence Interval for θ.

$\alpha = 0.05 \qquad \alpha/2 = 0.025 \qquad Z_{1-\alpha/2} = Z_{0.975} = 1.96$

$$\left[\theta - Z_{1-\alpha/2}\sqrt{\frac{\theta^2}{n}} \,,\ \theta + Z_{1-\alpha/2}\sqrt{\frac{\theta^2}{n}}\right] \qquad \theta = \theta_{ML} = 3.86$$

$$\sqrt{\frac{\theta^2}{n}} = \sqrt{\frac{(3.86)^2}{10}} = 1.2206$$

with %95 confidence $\hat{\theta}$ is $\left[3.86 - 1.96 \times 1.2206 \,,\ 3.86 + 1.96 \times 1.2206\right] = \left[1.467 \,,\ 6.2524\right]$

(L) explain if you can use the asymptotic distribution of $\hat{\theta}$ to obtain an approximate %95 confidence Interval for μ, and if possible, obtain such interval.

Subject .... Date ....

It is claimed that a new diet will reduce a person's weight by more than 4.5 Kilograms on average in a period of 2 weeks. the weights of 7 women who followed this diet were recorded before and after the 2 week period.

	$\boxed{1}$	$\boxed{2}$	$\boxed{3}$	$\boxed{4}$	$\boxed{5}$	$\boxed{6}$	$\boxed{7}$
Before	58.6	60.3	61.7	69.0	64.0	62.6	56.7
after	60.0	54.9	58.1	62.1	58.5	59.9	54.4

Can we state at a 1% significant level that the claim about the diet is false?

$\boxed{1}$	$\boxed{2}$	$\boxed{3}$	$\boxed{4}$	$\boxed{5}$	$\boxed{6}$	$\boxed{7}$
-1.4	+5.4	+3.6	+6.9	+5.5	+2.7	+2.3

$\bar{X} = \text{mean(before - after)} = 3.557143$ $\quad \mu_0 = 4.5$

$$SE = \sqrt{\frac{\text{var(before - after)}}{n}} = \sqrt{\frac{7.70619}{7}} = 1.04923$$

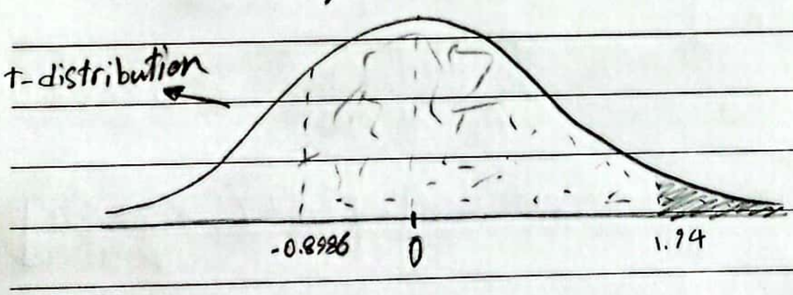
n is small $\sim$ we use t-distribution $\quad \sigma$ is unknown

D = Before - after

H_0: $D \leq 4.5$

H_1: $D > 4.5$

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{3.557143 - 4.5}{1.04923} = -0.8986$$



P-value > 0.05

=> there is significant error for test

Subject .... Date ....

Suppose that X is a discrete Random variable with the following probability mass function, where $0 \leq \theta \leq 1$ is a parameter.

X	0	1	2	3
$P(X=x)$	$2\theta/3$	$\theta/3$	$2(1-\theta)/3$	$(1-\theta)/3$

(a) compute expected value of X.

$$E[X] = \sum_{i=1}^{n} x f_x(x) = 0 \times 2\theta/3 + 1 \times \theta/3 + 2 \times \frac{2(1-\theta)}{3} + 3 \times \frac{(1-\theta)}{3}$$

$$= 0 + \frac{\theta}{3} + \frac{4}{3} - \frac{4}{3}\theta + \frac{3}{3} - \frac{3\theta}{3} = \frac{7}{3} - 2\theta$$

(b) the following 10 independent observations were take from such a distribution:

3, 0, 2, 1, 3, 2, 1, 0, 2, 1

calculate the Maximum likehood ~~distribution~~ estimate $\hat{\theta}$ of θ.

$$L(\theta) = \prod_{i=1}^{n} P(x_i, \theta) = \left(\frac{1-\theta}{3}\right)^2 \times \left(\frac{2\theta}{3}\right)^2 \times \left(\frac{2(1-\theta)}{3}\right)^3 \times \left(\frac{\theta}{3}\right)^3 = \frac{(1-\theta)^5 \theta^5 2^5}{3^{10}}$$

log. likehood function: $\ell(\theta)$:

$$\ell(\theta) = 5\log(1-\theta) + 5\log(\theta) + 5\log 2 - 10\log 3$$

Score function $\ell_*(\theta) = \frac{d}{d\theta}(\ell(\theta)) = 5\frac{-1}{1-\theta} + 5\frac{1}{\theta}$

Maximum likehood : $\ell_*(\theta) = 0$

$$\Rightarrow 5\frac{-1}{1-\theta} + 5\frac{1}{\theta} = 0 \longrightarrow \frac{1}{1-\theta} = \frac{1}{\theta} \longrightarrow 2\theta = 1 \Rightarrow \boxed{\theta = 1/2}$$

(c) calculate the asymptotic distribution of $\hat{\theta}$.

$$\boxed{J(\theta) = -\frac{d^2}{d\theta^2}(\ell(\theta)) = -\left[\frac{-5}{(1-\theta)^2} + 5\frac{-1}{\theta^2}\right] = \frac{5}{(1-\theta)^2} + \frac{5}{\theta^2}}$$

5 clips TM

Subject .... Date ....

$$J^{-1}(\theta) = \frac{1}{\frac{5}{(1-\theta)^2} + \frac{5}{\theta^2}} \qquad \hat{\theta} \sim N(\theta, J^{-1}(\theta))$$

$$\hat{\theta} \sim N\left(\theta, \frac{1}{\frac{5}{(1-\theta)^2} + \frac{5}{\theta^2}}\right)$$

(d) give a %95 (approximate) confidence interval for θ.

$$\alpha = 0.05 \qquad \alpha/2 = 0.025 \qquad Z_{1-\alpha/2} = Z_{0.975} = 1.96$$

$$\text{interval}: \left[\theta - Z_{1-\frac{\alpha}{2}}\left(\frac{1}{\frac{5}{(1-\theta)^2} + \frac{5}{\theta^2}}\right), \theta + Z_{1-\frac{\alpha}{2}}\left(\frac{1}{\frac{5}{(1-\theta)^2} + \frac{5}{\theta^2}}\right)\right]$$

$$\theta = \theta_{ML} = \frac{1}{2}$$

$$\Rightarrow \left[\frac{1}{2} - 1.96\left(\frac{1}{20+20}\right), \frac{1}{2} + 1.96\left(\frac{1}{20+20}\right)\right]$$

$$\Rightarrow \left[\frac{1}{2} - 0.049, \frac{1}{2} + 0.049\right]$$

$$\Longrightarrow [0.451, 0.549]$$

2- Researchers want to examine the effect of perceived control on health complaints of geriatric patients in a long-term care facility. thirty patients are randomly selected to participate in the study. Half are given a plant to care for and half are given a plant but the care is conducted by the staff. Number of Health complaints are recorded for each patient over the following seven days.

Control (X)	23	12	6	15	18	5	21	18	34	10	14	19	23	23	8
No control (Y)	35	21	26	24	17	~~22~~ 23	37	22	16	38	41	27	23	24	32

(a) what is the null Hypothesis in this study?

$$H_0: \mu_x - \mu_y = 0 \qquad H_1: \mu_x - \mu_y < 0$$

$\bar{X} = 16.6 \qquad \bar{Y} = 27.06 \qquad n_X = 15$

$S_X^2 = 60.68 \qquad S_Y^2 = 59.92 \qquad n_Y = 15$

(b) what is the alternative Hypothesis?

$$H_1 : \mu_X - \mu_Y < 0$$

(C) give an appropriate statistical test and assumptions underling its use.

$$t = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{S_X^2/n_X + S_Y^2/n_Y}}$$

(d) compute the value of statistical test on the observed data

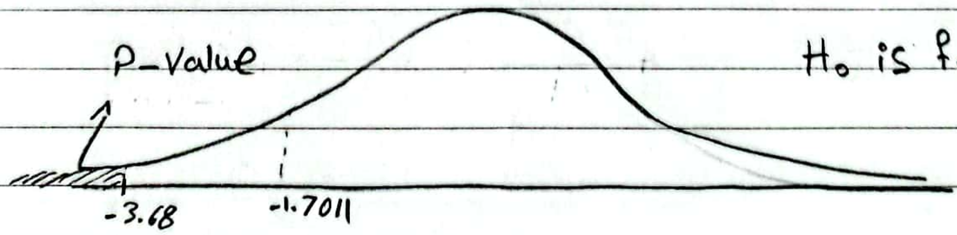
$$T = \frac{16.6 - 27.06 - 0}{\sqrt{60.68/15 + \frac{59.92}{15}}} = \frac{-10.46}{2.835} = -3.6895$$

(e) what are the degree of freedom?

$$\nu = \frac{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)^2}{\frac{(S_X^2/n_X)^2}{n_X - 1} + \frac{(S_Y^2/n_Y)^2}{n_Y - 1}} = \frac{(8.04)^2}{\frac{16.36 + 15.95}{14}} = \frac{64.64}{32.31} \times 14$$

$\Rightarrow \boxed{\nu \sim 28}$

(f) IS there a significant difference between two groups? interpret your answer.



H_0 is false

Subject .... Date ....

(f) Is there a significant difference between the two groups? Interpret your answer

Yes, the evidence againt H_0 is pretty strong. It would be prudent to use a perceived control to reduce the number of complaint.

(g) ~~is there assign~~ Is your answer feasible of a causal interpretation? explain.

(h) if you have made an error, would it be a type I or a Type II error? Explain your answer.

Type I is rejecting the null Hypothesis when it is true and type II error is accepting null Hypothesis when it is incorrect. In this Problem since we reject the null Hypothesis, we may have a type I error.

Subject .... Date ....

Assume that $X_1, X_2, \dots X_n$ are independent realizations of a random variable $X \sim N(0, 6^2)$

a) write the density function of X ?

$$P_x(\mu, 6^2) = \frac{1}{\sqrt{2\Pi 6^2}} \exp(-(x-\mu)^2/26^2)$$

$$P_x(1, 6^2) = \frac{1}{\sqrt{2\Pi 6^2}} \exp(-(x-1)^2/26^2)$$

(b) find the likehood function $L(6^2)$.

$$L(6^2) = \prod_{i=1}^{n} P(x_i, 6^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\Pi}} (6^2)^{-\frac{1}{2}} \exp(-(x-1)^2/26^2)$$

(C) obtain the log-likehood function $l(6^2)$.

$$l(6^2) = \log(L(6^2)) = -\frac{n}{2}\log 2\Pi - \frac{n}{2}\log 6^2 - \sum_{i=1}^{n} \frac{1}{26^2}(x-1)^2$$

$$= -\frac{n}{2}\log 2\Pi - \frac{n}{2}\log 6^2 - \frac{1}{26^2}\sum_{i=1}^{n}(x-1)^2$$

(d) obtain the score function $l_*(6^2)$

$$l_*(6^2) = \frac{d}{d6^2}(l(6^2)) = \frac{-n}{26^2} + \frac{1}{26^4}\sum_{i=1}^{n}(x-1)^2$$

(e) compute the Maximum likehood estimate $\hat{6}^2$ of 6^2.

$$l_*(6^2) = 0 \Rightarrow \frac{-n}{26^2} + \frac{1}{26^4}\sum_{i=1}^{n}(x-1)^2 \Rightarrow \boxed{6^2 = \frac{\sum_{i=1}^{n}(x-1)^2}{n}}$$

Subject .... Date ....

(f) check whether $\hat{\sigma}^2$ is a biased or an unbiased estimator of σ^2.

$B(\hat{\sigma}^2) = E[\hat{\sigma}^2] - \sigma^2$

(i) $E[\hat{\sigma}^2] = E\left[\frac{\sum_{i=1}^{n}(X-1)^2}{n}\right] = \frac{1}{n}\sum_{i=1}^{n} E[(X_i-1)^2] \xrightarrow{iid} \frac{n}{n} E[(X-1)^2]$

(ii) $X \sim N(1, \sigma^2) \longrightarrow X-1 \sim N(0, \sigma^2)$

(i), (ii) $\Rightarrow E[(X-1)^2] \xrightarrow{E(X)=0} E[(X-1)^2] = Var(X-1) = \sigma^2$

$\Rightarrow \hat{\sigma}^2$ is an unbiased estimator

Another way to solve:

$E[X^2 - 2X + 1] = E[X^2] - 2E[X] + 1 = E[X^2] - 2 + 1 = E[X^2] - 1$

$= E[X^2] - E[X]^2 = Var[X] = \sigma$

(g) obtain an asymptotic approximation for the distribution of $\hat{\sigma}^2$.

$\hat{\sigma}^2 \sim N(\sigma^2, I(\hat{\sigma}^2)^{-1})$

$I(\sigma^2) = -\frac{d(S_n(\sigma^2))}{d\sigma^2} = +\frac{n}{2\sigma^4} - \frac{1}{(\sigma^2)^3}\sum_{i=1}^{n}(X-1)^2$

from (e) we have $n\sigma^2 = \sum_{i=1}^{n}(X-1)^2$ $\longrightarrow + \frac{n}{2\sigma^4} - \frac{n\sigma^2}{(\sigma^2)^3} = \frac{n}{2\sigma^4} - \frac{2n}{2\sigma^4} = \boxed{-\frac{n}{2\sigma^4}}$

$\Rightarrow I(\hat{\sigma}^2)^{-1} = \boxed{+\frac{2\sigma^4}{n}}$ $\quad \hat{\sigma}^2 \sim N\left(\sigma^2, \frac{2\sigma^4}{n}\right)$

(h) obtain the exact distribution of $\hat{\sigma}^2$

$\hat{\sigma}^2 = \frac{\sum(X-1)^2}{n} = \frac{\sigma^2}{n}\sum_{i=1}^{n}\left(\frac{X-1}{\sigma}\right)^2$ $\quad \Big| \quad \frac{X-1}{\sigma} \sim N(0,1)$

$\left(\frac{X-1}{\sigma}\right)^2 \sim \chi^2_n \longrightarrow$ chi-square

clips

Subject .... Date ....

(h) cont.

$$\Rightarrow \hat{\sigma}^2 \sim \frac{\sigma^2}{n} \chi^2_n$$

(i) compute $\hat{\sigma}^2$ with data (see R script below)

1	2	3	4	5	6	7	8	9	10
2.84	2.66	-0.62	-1.82	2.18	2.10	2.84	-0.98	1.12	-0.89

$$\hat{\sigma}^2 = \frac{\Sigma(x-1)^2}{n} = \frac{>\text{sum}((x-1)^2)}{>\text{length}(x)} = \frac{30.1566}{10} = 3.01566$$

(j) Use the asymptotic distribution of $\hat{\sigma}^2$ to compute an approximation of %95 confidence Interval for it.

$$\%95 \longrightarrow \alpha = 0.05 \longrightarrow \boxed{z_{\alpha/2} = 1.96}$$

Asymptotic SE $= \sqrt{\frac{2\hat{\sigma}^4}{n}} = 1.3486$ From above we had $\boxed{\hat{\sigma}^2 = 3.01566}$

$$\Rightarrow \left[\hat{\sigma}^2 - z_{\alpha/2} * SE,\ \hat{\sigma}^2 + z_{\alpha/2} * SE\right]$$

$$\Rightarrow [3.01566 - 1.96 * 1.3486,\ 3.01566 + 1.96 * 1.3486]$$

$$\rightarrow [0.375, 5.65] \ \%95 \text{ confidence}$$

(K) Use exact distribution of $\hat{\sigma}^2$ to compute a %95 confidence Interval.

$$\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_n \Rightarrow P\left(\chi^2_{\frac{\alpha}{2},n} \leqslant \frac{n\hat{\sigma}^2}{\sigma^2} \leqslant \chi_{1-\frac{\alpha}{2},n}\right) = 0.95$$

$$\Rightarrow P\left(3.25 \leqslant \frac{n\hat{\sigma}^2}{\sigma^2} < 20.48\right) = P\left(\frac{n\hat{\sigma}^2}{20.48} \leqslant \sigma^2 \leqslant \frac{n\hat{\sigma}^2}{3.25}\right)$$

$n = 10$

$\hat{\sigma}^2 = 3.01566 \Rightarrow P(1.4724 \leqslant \sigma^2 \leqslant 9.2789) = 0.95$

Subject .... Date ....

For a certain disease, there exists a treatment that cures 70% of cases. A laboratory proposes a new treatment claiming that it is better than pervious one. out of 200 patients received new treatments, 148 of them have been cured. As the expert in charge of deciding whether the new treatment shoud be authorized? what are your conclusions?

$\boxed{1}$ $X \sim Bin(\Pi)$

$\boxed{2}$ $X_1, X_2, X_3, \ldots X_n$ iid (sample)

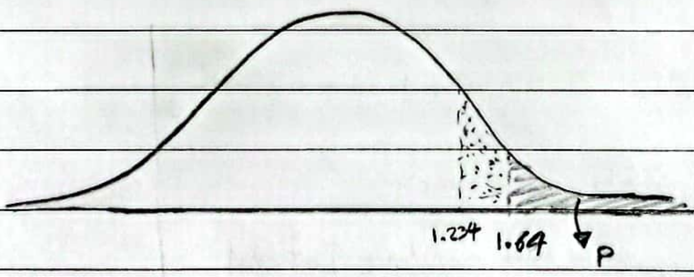
$\boxed{3}$ $H_0: \Pi \leqslant 0.7$

$H_1: \Pi > 0.7$

definition of statistics: $\frac{\Pi - \Pi_0}{SE_0} = z \Rightarrow z = \frac{0.74 - 0.7}{0.0324} = 1.234$

$\Pi = \frac{148}{200} = 0.74$ $\Pi_0 = 0.7$

$SE = \sqrt{\frac{\Pi_0(1-\Pi_0)}{n}} = \sqrt{\frac{0.7(1-0.7)}{200}} = 0.0324$



n = 200 is large enough so we use normal distribution

From above plot: P-Value > 0.05 $\longrightarrow$ there is a significant error for test

clips

Subject .... Date ....

Assume that $y_1, y_2, \dots, y_n$ are independent realizations of positive continuous random variable Y with pdf:

$$f(y,\theta) = \theta(1+\theta)\, y^{\theta-1}(1-y) \qquad \theta>0,\ y\in(0,1)$$

(a) show that $f(y,\theta)$ is a density function.

$$\int_0^1 f(y,\theta)\,dy = \int_0^1 \theta(1+\theta)\, y^{\theta-1}(1-y)\,dy = \theta(1+\theta)\left[\int_0^1 y^{\theta-1}\,dy - \int_0^1 y^{\theta}\,dy\right]$$

$$= \theta(1+\theta)\left[\frac{1}{\theta}\, y^{\theta}\right]_0^1 - \theta(1+\theta)\left[\frac{1}{\theta+1}\, y^{\theta+1}\right]_0^1 = 1+\theta-\theta = \boxed{1}$$

(b) obtain the likehood function, $L(\theta)$, and log likehood function $\ell(\theta)$.

$$L(\theta) = \prod_{i=1}^{n} \theta(1+\theta)\, y_i^{\theta-1}(1-y_i)$$

$$\ell(\theta) = n\log\theta + n\log(1+\theta) + (\theta-1)\sum_{i=1}^{n}\log y_i + \sum_{i=1}^{n}\log(1-y_i)$$

(c) write the score function for θ and verify that with the data

$y = ($ ① 0.77, ② 0.95, ③ 0.62, ④ 0.85, ⑤ 0.27, ⑥ 0.01, ⑦ 0.29, ⑧ 0.67, ⑨ 0.80, ⑩ 0.38, ⑪ 0.73, ⑫ 0.18, ⑬ $0.13)$

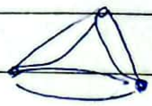
$\hat{\theta} = 1.513$ is the maximum likehood estimate of θ $\left(\sum_{i=1}^{n} \ln y_i = -13.765\right)$

$$\ell_*(\theta) = \frac{n}{\theta} + \frac{n}{1+\theta} + \sum_{i=1}^{n}\log y_i \quad \Rightarrow \quad \ell_*(\theta) = 0$$

$$\Rightarrow \frac{n(2\theta+1)}{\theta(1+\theta)} = -\sum_{i=1}^{n}\log y_i \rightarrow 2n\theta + n = -\theta\sum_{i=1}^{n}\log y_i - \theta^2\sum_{i=1}^{n}\log y_i$$

$$\Rightarrow \theta^2\sum_{i=1}^{n}\log y_i + \theta\left(2n + \sum_{i=1}^{n}\log y_i\right) + n = 0 \qquad (ax^2+bx+c=0)$$

$a = \sum_{i=1}^{n}\log y_i \qquad b = 2n + \sum_{i=1}^{n}\log y_i \qquad c = n$



Subject .... Date ....

$$\Delta = b^2 - 4ac = \left(2n + \sum_{i=1}^{n} \log y_i\right)^2 - 4\left(\sum_{i=1}^{n} \log y_i\right)(n)$$

$$\hat{\theta}_{ML} = \frac{-b \pm \sqrt{\Delta}}{2a} = \frac{-\left(2n + \sum_{i=1}^{n} \log y_i\right) \pm \sqrt{\Delta}}{2\left(\sum_{i=1}^{n} \log y_i\right)}$$

$n = 13$, $\sum_{i=1}^{n} \log y_i = -13.765$

$$\Delta = \left(2 \times 13 + (-13.765)\right)^2 - 4(-13.765)(13)$$

$$\Delta = (26 - 13.765)^2 - 52(-13.765) = 149.695 + 715.78 = \boxed{865.475}$$

$$\sqrt{\Delta} = \boxed{29.418}$$

$$b = 2n + \sum_{i=1}^{n} \log y_i = 26 + (-13.765) = \boxed{12.235}$$

$$a = \sum_{i=1}^{n} \log y_i = -13.765$$

$$\Rightarrow \hat{\theta}_{ML} = \frac{-12.235 \pm 29.418}{2 \times (-13.765)} \Rightarrow \hat{\theta}_{ML} = 1.513 \text{ or } -0.624$$

Since $\theta > 0 \Rightarrow \boxed{\theta_{ML} = 1.513}$

(d) obtain an approximation for the distribution of $\hat{\theta}$. build a confidence interval for θ of approximate level 0.95.

$$\hat{\theta} \sim N(\theta, J^{-1}(\theta))$$

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

$$\Rightarrow \hat{\theta} \sim N(\theta, J^{-1}(\theta)) \longrightarrow \hat{\theta} \sim N\left(\theta, \frac{\theta^2(1+\theta)^2}{n(2\theta^2 + 2\theta + 1)}\right)$$

Subject .... Date ....

CI %95 : $\alpha = 0.05$ $\quad \alpha/2 = 0.025$ $\quad Z_{\alpha/2} = 1.96$

$$CI = \left[\theta - Z_{\alpha/2}\sqrt{\frac{\theta^2(1+\theta)^2}{n(2\theta^2+2\theta+1)}},\ \theta + Z_{\alpha/2}\sqrt{\frac{\theta^2(1+\theta)^2}{n(2\theta^2+2\theta+1)}}\right]$$

$n = 13,\ \theta_{ML} = 1.513$, $\dfrac{\theta^2(1+\theta)^2}{n(2\theta^2+2\theta+1)} = \dfrac{(1.513)^2(2.513)^2}{13(2\times(1.513)^2 + 2\times 1.513 + 1)}$

$$= \frac{2.289 \times 5.076}{13(4.578 + 3.026 + 1)} = \frac{11.618}{111.852} = \sqrt{0.1038}$$

$\Rightarrow$ CI : $[1.513 - 1.96 \times 0.3223,\ 1.513 + 1.96 \times 0.3223]$

CI : $[1.513 - 0.631,\ 1.513 + 0.631]$

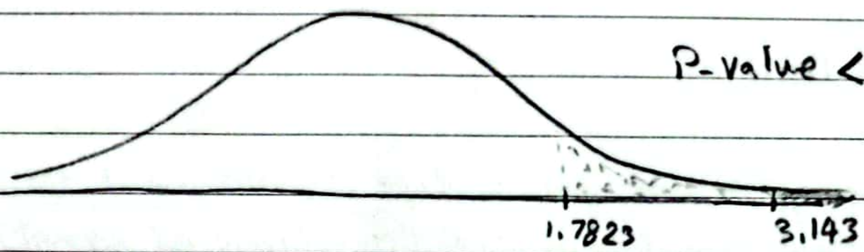
CI : $[0.882,\ 2.144]$

(e) Test $H_0 : \theta = \frac{1}{2}$ vs $H_1 : \theta > 1/2$ at significant level 0.05.

statistics : $\dfrac{\theta - \theta_0}{\sigma} \rightarrow \dfrac{1.513 - 0.5}{0.3223} = \dfrac{1.013}{0.3223} = 3.143$

$\rightarrow$ T-distribution with n-1 degree

$T_{12,\ 0.05} = 1.7823$



P-value < 0.05

H_0 is False.

Subject .... Date ....

2 - consider the following $n=10$ iid observation from $X \sim N(\mu, \sigma^2)$

2.38 3.04 6.62 3.64 3.76 6.93 4.42 0.97 2.13 2.61

(a) obtain a %80 confidence interval for σ^2.

$$Q = \frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad \alpha = 0.2 \quad \frac{\alpha}{2} = 0.1$$

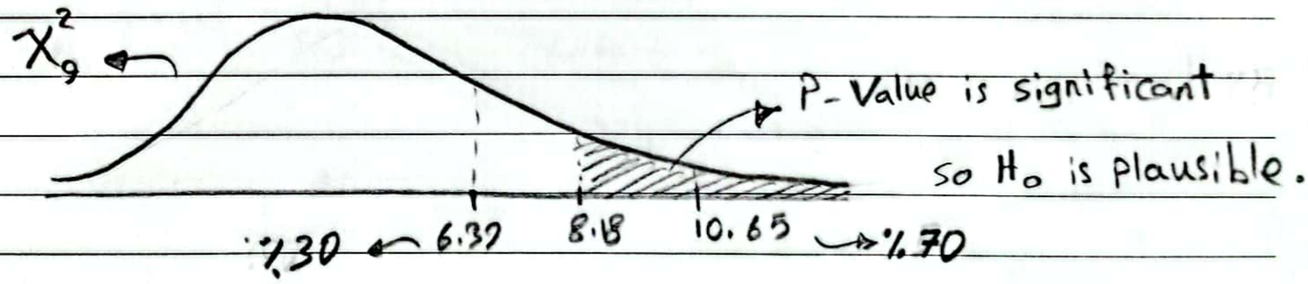
$$CI: \left[\chi^2_{1-\frac{\alpha}{2}, n-1} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{\frac{\alpha}{2}, n-1}\right]$$

$$CI: \left[\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}}\right]$$

$$CI: \left[\frac{9 \times 3.6382}{14.68}, \frac{9 \times 3.6382}{4.1681}\right] = [2.23, 7.8558]$$

(b) Test the Hypothesis $H_0: \sigma^2 = 4$ against the alternative $H_1: \sigma^2 > 4$

$$Q = \frac{(n-1)S^2}{\sigma^2} = \frac{9 \times 3.6382}{4} = 8.1859$$



3. True or False?

(i) if a sample size is large, then the shape of a histogram of the data will be approximately normal, even if population distribution is not normal.

False, if $n \to \infty$ → Sample is population so above sentence is incorrect

Subject .... Date ....

iii) if a sample size is large, the the shape of the sampling distribution of the sample mean will be approximately normal, even if the population distribution is not normal.

True, Based on central limit theorem, when the sample size is large enough, the sample mean distribution is approximately normal, and no matter what population the sample was drawn from.

(b) In a study of the effects of smoking during Pregnancy, measurments on basis of mothers who smokes during pregnancy were compared to measurments on babies of mothers who did not. A %95 confidence interval for the difference in mean weight [-344.9548, 159.9548] gerams. what can be said from this statement about P-value for Hypothesis that mean difference is zero?

$H_0: \mu_1 - \mu_2 \neq 0$ $\quad$ $\bar{X}$: smoke $\quad$ $\bar{Y}$: donot smoke

$H_1: \mu_1 - \mu_2 = 0$ $\quad$ $\alpha = 0.05$ $\quad$ $\alpha/2 = 0.025$ $\quad$ $Z_{\alpha/2} = 1.96$

$$CI: \left[(\bar{X}-\bar{Y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} \, , \, (\bar{X}-\bar{Y}) + Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}\right]$$

$$\text{statistics:} \quad \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$$

$$[-344.9548, 159.9548] \longrightarrow \frac{159.9548 - (-344.9548)}{2} = Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$$

$$\Rightarrow \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} = 128.8034$$

$$\Rightarrow (\bar{X}-\bar{Y}) + Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} = 159.9548 \longrightarrow \boxed{\bar{X} - \bar{Y} = -92.5}$$

$$\Rightarrow \text{statistics} = \frac{-92.5}{128.8034} = -0.7181$$

$\longrightarrow$ Sum of area in two side of normal distribution is large so H_0 is plausible.

Subject .... Date ....

(C) consider Standard notation used in the course. what is the difference between:

(a) the mean of Y and mean of $\bar{Y}$?

(b) the standard deviation of Y and the standard deviation of $\bar{Y}$?

(c) the standard deviation of Y and Standard error of $\bar{Y}$? Does any relation exist between two quantities in each statement?