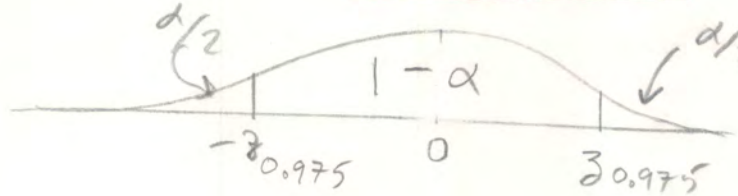
ASSIGNMENT 3

① (a)



0.95

(b) $z_{0.975} = 1.96$

(c) $P(-z_{0.995} < Z < z_{0.995}) = P(z_{.005} < Z < z_{.995}) = 0.99$

(d) $z_{0.995} = (2.57 + 2.58)/2 = 2.575$

(e) $1 - \alpha$

② By the (modified) Central Limit Theorem,

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} \xrightarrow{d} Z \sim N(0,1)$$

$$1 - \alpha = P(-z_{1-\alpha/2} < Z < z_{1-\alpha/2}) \approx P(-z_{1-\alpha/2} < Z_n < z_{1-\alpha/2})$$

$$= P\left(-z_{1-\alpha/2} < \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} < z_{1-\alpha/2}\right)$$

$$= P\left(-z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}} < \bar{X}_n - \mu < z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}}\right)$$

$$= P\left(-\bar{X}_n - z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}} < -\mu < -\bar{X}_n + z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}}\right)$$

$$= P\left(\bar{X}_n + z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}} > \mu > \bar{X}_n - z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}}\right)$$

$$= P\Bigg(\underbrace{\bar{X}_n - z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}}}_{\text{Lower confidence Limit}} < \mu < \underbrace{\bar{X}_n + z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}}}_{\text{Upper confidence Limit}}\Bigg)$$

③ (a) $\hat{\lambda} = \bar{X}_n = 9.2$

Since variance of Poisson $= \lambda = \sigma^2$, take $\hat{\sigma}_n^2 = \bar{X}_n$, consistent by Law of Large Numbers.

By Question 2, 95% CI is $\bar{X}_n \pm z_{0.975}\sqrt{\frac{\bar{X}_n}{n}}$

$$= 9.2 \pm 1.96\sqrt{9.2/30} = 9.2 \pm 1.09$$

$$= (8.11, 10.29)$$

(b) Yes, since 8 is outside the estimated range of values. Of course this depends on α.

④ (a) $\hat{\theta} = \bar{X}_n = 0.6$, $\hat{\sigma}_n^2 = \bar{X}_n(1-\bar{X}_n) \xrightarrow{P} \theta(1-\theta) = \sigma^2$ by LLN & continuous mappings, so 95% CI is

$$\bar{X}_n \pm z_{1-\alpha/2}\frac{\hat{\sigma}_n}{\sqrt{n}} = 0.6 \pm 1.96\,\frac{\sqrt{0.6(1-0.6)}}{\sqrt{100}}$$

$$= 0.6 \pm 0.096 = (0.504, 0.696)$$

But we were asked for a percentage, so it's

$$60\% \pm 9.6\% = (50.4\%, 69.6\%)$$

(4 b) (i) Want $z_{1-\alpha/2} \frac{\hat{\sigma}_n}{\sqrt{n}} = 0.03$

$$1.96\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}} \approx 1.96\sqrt{\frac{\theta(1-\theta)}{n}}$$

$$= 1.96\sqrt{\frac{0.6 * 0.4}{n}} \overset{\text{set}}{=} 0.03$$

Solve for n

$$\Rightarrow \sqrt{n} = \frac{1.96\sqrt{0.6*0.4}}{0.03}$$

$$\Rightarrow n = \frac{1.96^2(0.6*0.4)}{0.03^2} = 1024.427$$

So take $n = 1025$

(ii) If true $\theta = 0.7$, $n = \frac{1.96^2(.7*.3)}{0.0009} = 896.37$

So take $n = 897$

(iii) If true $\theta = 0.9$, $n = \frac{1.96^2(.9*.1)}{0.0009}$
$= 384.16$

(iv) If true $\theta = 0.3$, $n = \frac{1.96^2(.3*.7)}{0.0009} = 896.37$ AGAIN

So $n = 897$

(4 b v.) Seek maximum value of $n = \frac{1.96^2 \theta(1-\theta)}{.03^2}$

Find θ yielding maximum value of $n = n(\theta)$

$$\frac{dn}{d\theta} = \frac{(1.96)^2}{.03^2} \frac{d}{d\theta}(\theta - \theta^2) = \frac{1.96^2}{.03^2}(1 - 2\theta)$$

$$\overset{set}{=} 0 \Rightarrow 2\theta = 1 \Rightarrow \theta = \frac{1}{2}$$

2nd derivative test;

$$\frac{d^2 n}{d\theta^2} = \frac{1.96^2}{.03^2}(-2) \text{ negative concave down } \cap$$

max

So the maximum required n occurs at $\theta = \frac{1}{2}$

And $$n = \frac{1.96^2 \left(\frac{1}{2} \cdot \frac{1}{2}\right)}{0.0009} = 1067.11$$

So take $n = 1{,}068$

A lot of participants for a little taste test, but typical of political polls.

(5) $\bar{x}_n \pm z_{1-\alpha/2} \frac{\hat{\sigma}_n}{\sqrt{n}} = 105 \pm 1.96 \frac{\sqrt{256}}{\sqrt{64}}$

$= 105 \pm 1.96 \frac{16}{8} = 105 \pm 3.92 = (101.08, 108.92)$

(6) (a) $\bar{x} = 8.96$ estimates $E(X_i) = \frac{1}{\lambda} = \mu$

(b) (i) $M_{\bar{X}}(t) = M_{\frac{1}{n}\sum_{i=1}^n X_i}(t) = M_{\sum_{i=1}^n X_i}\left(\frac{t}{n}\right)$

$\overset{ind}{=} \prod_{i=1}^n M_{X_i}(t/n) = \prod_{i=1}^n \left(1 - \frac{t}{n\lambda}\right)^{-1} = \left(1 - \frac{t}{n\lambda}\right)^{-n}$

MGF of Gamma$(\alpha = n, \lambda' = n\lambda)$

(ii) $M_Y(t) = M_{2\lambda n \bar{X}}(t) = M_{\bar{X}}(2\lambda n t)$

$= \left(1 - \frac{2\cancel{\lambda n} t}{\cancel{\lambda n}}\right)^{-n} = (1 - 2t)^{-\frac{2n}{2}}$

MGF of $\chi^2(\gamma = 2n)$

(iii) $1 - \alpha = P(\chi^2_{0.025} < Y < \chi^2_{0.975})$

$= P(\chi^2_{0.025} < 2\lambda n \bar{X} < \chi^2_{0.975})$

$= P\left(\frac{\chi^2_{0.025}}{2n\bar{X}} < \lambda < \frac{\chi^2_{0.975}}{2n\bar{X}}\right)$

$= P\left(\frac{2n\bar{X}}{\chi^2_{0.025}} > \frac{1}{\lambda} > \frac{2n\bar{X}}{\chi^2_{0.975}}\right)$

$= P\left(\frac{2n\bar{X}}{\chi^2_{0.975}} < \mu < \frac{2n\bar{X}}{\chi^2_{0.025}}\right) = P(L < \mu < U)$

(6biv) $\left(\frac{2n\bar{x}}{\chi^2_{0.975}}, \frac{2n\bar{x}}{\chi^2_{0.025}} \right) =$

$= \left(\frac{(2)(14)(8.96)}{44.46}, \frac{(2)(14)(8.96)}{15.31} \right) = (5.64, 16.39)$

(6c) (i) CI for μ: $Var(X_1) = \sigma^2 = \frac{1}{\lambda^2}$. Taking $\hat{\lambda}_n = 1/\bar{X}_n$, have $\hat{\sigma}_n^2 = \frac{1}{\hat{\lambda}_n^2} = \bar{X}_n^2$, consistent by LLN and continuous mapping. So

$\bar{x}_n \pm z_{1-\alpha/2} \frac{\hat{\sigma}_n}{\sqrt{n}} = 8.96 \pm 1.96 \frac{8.96}{\sqrt{14}} = 8.96 \pm 4.69$

$= (4.27, 13.65)$

(ii) $0.95 \approx P\left(\bar{X}_n - 1.96 \frac{\bar{X}_n}{\sqrt{n}} < \mu < \bar{X}_n + 1.96 \frac{\bar{X}_n}{\sqrt{n}} \right)$

$= P\left(\frac{1}{\bar{X}_n - 1.96 \frac{\bar{X}_n}{\sqrt{n}}} > \frac{1}{\mu} > \frac{1}{\bar{X}_n + 1.96 \frac{\bar{X}_n}{\sqrt{n}}} \right)$

$= P\left(\frac{1}{\bar{X}_n + 1.96 \frac{\bar{X}_n}{\sqrt{n}}} < \lambda < \frac{1}{\bar{X}_n - 1.96 \frac{\bar{X}_n}{\sqrt{n}}} \right)$

So the numerical confidence interval is

$\left(\frac{1}{13.65}, \frac{1}{4.27} \right) = (0.073, 0.234)$

7 (a) $F_{T_n}(t) = \left(\frac{t}{\theta}\right)^n I(0<t<\theta) + I(t \geq \theta)$

(b)(i) With probability one, $0 < T_n < \theta \iff 0 < \frac{T_n}{\theta} < 1$

$\iff 0 > -\frac{T_n}{\theta} > -1 \iff 1 > 1 - \frac{T_n}{\theta} > 0$

$\iff 0 < n\left(1 - \frac{T_n}{\theta}\right) < n \iff 0 < Y_n < n.$

(ii) For $0<y<n$, $F_{Y_n}(y) = P(Y_n \leq y)$

$= P\left(n\left(1 - \frac{T_n}{\theta}\right) \leq y\right) = P\left(1 - \frac{T_n}{\theta} \leq \frac{y}{n}\right)$

$= P\left(T_n/\theta \geq 1 - \frac{y}{n}\right) = P\left(T_n \geq \theta\left(1 - \frac{y}{n}\right)\right)$

$= 1 - F_{T_n}\left(\theta\left(1 - \frac{y}{n}\right)\right) = 1 - \left[\frac{\theta\left(1 - \frac{y}{n}\right)}{\theta}\right]^n$

$= 1 - \left(1 - \frac{y}{n}\right)^n$, so

$$F_{Y_n}(y) = \left[1 - \left(1 - \frac{y}{n}\right)^n\right] I(0<y<n) + I(y \geq n)$$

(iii) For any $y>0$, eventually $y<n$ and

$$\lim_{n\to\infty} F_{Y_n}(y) = \lim_{n\to\infty}\left[1 - \left(1 - \frac{y}{n}\right)^n\right]$$

$$= 1 - \lim_{n\to\infty}\left(1 + \frac{-y}{n}\right)^n = 1 - e^{-y},$$

cumulative distribution function of standard exponential.

(7c) $1-\alpha = P(0 < Y < y_{1-\alpha})$

$$\approx P\left(0 < n\left(1 - \frac{T_n}{\theta}\right) < y_{1-\alpha}\right)$$

$$= P\left(0 < 1 - T_n/\theta < y_{1-\alpha}/n\right)$$

$$= P\left(0 > \frac{T_n}{\theta} - 1 > -\frac{y_{1-\alpha}}{n}\right) = P\left(1 > \frac{T_n}{\theta} > 1 - \frac{y_{1-\alpha}}{n}\right)$$

$$= P\left(1 < \frac{\theta}{T_n} < \frac{1}{1 - \frac{y_{\alpha-1}}{n}}\right)$$

$$= P\left(T_n < \theta < \frac{T_n}{\frac{n - y_{\alpha-1}}{n}}\right) = P\left(T_n < \theta < \frac{n}{n - y_{\alpha-1}} T_n\right)$$

So the $(1-\alpha)100\%$ confidence interval is

$$\left(T_n, \frac{n}{n - y_{1-\alpha}} T_n\right)$$

(d) No

(e) $1-\alpha = F_Y(y_{\alpha-1}) = 1 - e^{-y_{\alpha-1}}$

$$\iff \alpha = e^{-y_{\alpha-1}} \iff \ln(\alpha) = -y_{\alpha-1}$$

$$\iff y_{\alpha-1} = -\ln(\alpha)$$

(7f) (i) $-\ln(0.05) = 2.9957 \approx 3$, so CI is

$$\left(3.905, \left(\frac{30}{30-3}\right) 3.905\right) = \left(3.905, \left(\frac{30}{27}\right) 3.905\right)$$

$$= (3.905, 4.339)$$

(ii) $\mu = \frac{\theta}{2}$, $\sigma^2 = \frac{\theta^2}{12}$. Taking $\hat{\theta} = 2\bar{X}_n$,

$$\hat{\sigma}_n^2 = \frac{(2\bar{X}_n)^2}{12} = \frac{4\bar{X}_n^2}{12} = \frac{1}{3}\bar{X}_n^2 \xrightarrow{P} \sigma^2 \text{ by}$$

LLN and continuous mapping. So,

$$0.95 \approx P\left(\bar{X}_n - 1.96\,\frac{\bar{X}_n/\sqrt{3}}{\sqrt{n}} < \mu < \bar{X}_n + 1.96\,\frac{\bar{X}_n}{\sqrt{3n}}\right)$$

$$= P\left(2\bar{X}_n - 2(1.96)\frac{\bar{X}_n}{\sqrt{3n}} < 2\mu < 2\bar{X}_n + 2(1.96)\frac{\bar{X}_n}{\sqrt{3n}}\right)$$

$$= P\left(2\bar{X}_n - 3.92\frac{\bar{X}_n}{\sqrt{3n}} < \theta < 2\bar{X}_n + 3.92\frac{\bar{X}_n}{\sqrt{3n}}\right)$$

So with $\bar{x}_n = 1.785$ & $n = 30$, have

$$(3.57 - 0.738, 3.57 + 0.738) = (2.832, 4.308)$$

(iii) I like the first one more. They are both large-sample approximations, but the first one is narrower (0.434 vs 1.476), and the lower confidence limit is a certainty.

(8) Z_n in Q2, Y in Q6, Y_n in Q7.