STA500 Introduction to Probability and Statistics 2, autumn 2018.

Solution exercise set 6

Exercises from the book:

9.13/9.15

Measurements of hardness: $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma^2)$. Both μ and σ unknown.

Confidence interval for μ:

$$\hat{\mu} = \bar{X}$$

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

$$\begin{aligned} P(-t_{\alpha/2,n-1} \leq \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_{\alpha/2,n-1}) &= 1 - \alpha \\ P(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}) &= 1 - \alpha \end{aligned}$$

I.e. a $(1-\alpha)100\%$ confidence interval for μ is given by: $[\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}]$

Observed: $n = 12$, $\bar{x} = 48.50$, $s = 1.5$. $\alpha = 0.1 \quad \Rightarrow \quad t_{\alpha/2,n-1} = t_{0.05,11} = 1.796$. Inserted this gives the 90% confidence interval for μ:

$$[48.50 - 1.796\frac{1.5}{\sqrt{12}}, 48.50 + 1.796\frac{1.5}{\sqrt{12}}] = \underline{\underline{[47.72, 49.28]}}$$

9.76

Measurements of hardness: $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma^2)$.

Confidence interval for σ:

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Theorem 8.4/collection of formulas gives that: $(n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$

$$\begin{aligned} \Rightarrow \qquad P(\chi^2_{1-\alpha/2,n-1} \leq (n-1)\frac{S^2}{\sigma^2} \leq \chi^2_{\alpha/2,n-1}) &= 1 - \alpha \\ P\left(\frac{\chi^2_{1-\alpha/2,n-1}}{(n-1)S^2} \leq \frac{1}{\sigma^2} \leq \frac{\chi^2_{\alpha/2,n-1}}{(n-1)S^2}\right) &= 1 - \alpha \\ P\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}} \geq \sigma^2 \geq \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}\right) &= 1 - \alpha \\ P\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}} \leq \sigma \leq \sqrt{\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}}\right) &= 1 - \alpha \end{aligned}$$

Observed: $n = 12$ and $s = 1.5$.

Further $\alpha = 0.10 \;\Rightarrow\; \chi^2_{0.05,11} = 19.675$ and $\chi^2_{0.95,11} = 4.575$.

90% confidence interval for σ: $\left[\sqrt{\frac{11\cdot 1.5^2}{19.675}}, \sqrt{\frac{11\cdot 1.5^2}{4.575}}\right] = \underline{\underline{[1.12, 2.33]}}$.

Exercise 1 in MLE-note:

a)

$$
\begin{aligned}
L(\beta) &= = \prod_{i=1}^{n} \frac{1}{\beta} e^{-x_i/\beta} = \frac{1}{\beta^n} e^{-\sum_{i=1}^n x_i/\beta} \\
l(\beta) &= \ln L(\beta) = \ln(1) - \ln(\beta)^n - \sum_{i=1}^{n} x_i/\beta = -n\ln(\beta) - \frac{1}{\beta}\sum_{i=1}^{n} x_i \\
\frac{\partial l(\beta)}{\partial \beta} &= -\frac{n}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^{n} x_i = 0 \\
& n\beta = \sum_{i=1}^{n} x_i \qquad \Rightarrow \quad \beta = \frac{1}{n}\sum_{i=1}^{n} x_i
\end{aligned}
$$

I.e MLE becomes $\hat{\beta} = \underline{\underline{\frac{1}{n}\sum_{i=1}^{n} X_i}}$.

b) For the exponential distribution we have $\mathrm{E}(X) = \beta$ and $\mathrm{Var}(X) = \beta^2$ and we thus get:

$$
\begin{aligned}
\mathrm{E}(\hat{\beta}) &= \mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^{n} X_i\Big) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n}\beta = \frac{1}{n}n\beta = \underline{\underline{\beta}} \quad \text{unbiased!} \\
\mathrm{Var}(\hat{\beta}) &= \mathrm{Var}\Big(\frac{1}{n}\sum_{i=1}^{n} X_i\Big) = \frac{1}{n^2}\mathrm{Var}\Big(\sum_{i=1}^{n} X_i\Big) \\
&\overset{indep.}{=} \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\beta^2 = \frac{1}{n^2}n\beta^2 = \underline{\underline{\frac{\beta^2}{n}}}
\end{aligned}
$$

c) We start by finding the second derivative of the log-likelihood at $\hat{\beta}$:

$$
\begin{aligned}
\frac{\partial l(\beta)}{\partial \beta} &= -\frac{n}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^{n} x_i \\
J(\beta) &= \frac{\partial^2 l(\beta)}{\partial \beta^2} = \frac{n}{\beta^2} - \frac{2}{\beta^3}\sum_{i=1}^{n} x_i \\
J(\hat{\beta}) &= \frac{n}{\hat{\beta}^2} - \frac{2n}{\hat{\beta}^3}\frac{\sum_{i=1}^{n} x_i}{n} = \frac{n}{\hat{\beta}^2} - \frac{2n}{\hat{\beta}^2} = -\frac{n}{\hat{\beta}^2}
\end{aligned}
$$

The Wald confidence interval then becomes:

$$
[\hat{\beta} - z_{\alpha/2}\sqrt{-1/J(\hat{\beta})}, \hat{\beta} + z_{\alpha/2}\sqrt{-1/J(\hat{\beta})}] = \underline{\underline{[\hat{\beta} - z_{\alpha/2}\sqrt{\hat{\beta}^2/n}, \hat{\beta} + z_{\alpha/2}\sqrt{\hat{\beta}^2/n}]}}
$$

d) With $n = 5$, $\hat{\beta} = 1274$, $z_{\alpha/2} = z_{0.025} = 1.96$, $\chi^2_{1-\alpha/2,2n} = \chi^2_{0.975,10} = 3.247$ and $\chi^2_{\alpha/2,2n} = \chi^2_{0.025,10} = 20.483$ we get the Wald interval

$$[\hat{\beta} - z_{\alpha/2}\sqrt{\hat{\beta}^2/n}, \hat{\beta} + z_{\alpha/2}\sqrt{\hat{\beta}^2/n}] = [1274 - 1.96\sqrt{1274^2/5},\ \ 1274 + 1.96\sqrt{1274^2/5}] = \underline{\underline{[157,\ \ 2391]}}$$

and the exact interval

$$[\frac{2n\hat{\beta}}{\chi^2_{\alpha/2,2n}}, \frac{2n\hat{\beta}}{\chi^2_{1-\alpha/2,2n}}] = [\frac{2 \cdot 5 \cdot 1274}{20.483}, \ \ \frac{2 \cdot 5 \cdot 1274}{3.247}] = \underline{\underline{[622,\ \ 3924]}}.$$

e) With $n = 50$, $\hat{\beta} = 1274$, $z_{0.025} = 1.96$, $\chi^2_{1-\alpha/2,2n} = \chi^2_{0.975,100} = 72.222$ and $\chi^2_{\alpha/2,2n} = \chi^2_{0.025,100} = 129.561$ we get the Wald interval

$$[\hat{\beta} - z_{\alpha/2}\sqrt{\hat{\beta}^2/n}, \hat{\beta} + z_{\alpha/2}\sqrt{\hat{\beta}^2/n}] = [1274 - 1.96\sqrt{1274^2/50},\ \ 1274 + 1.96\sqrt{1274^2/50}] = \underline{\underline{[921,\ \ 1627]}}$$

and the exact interval

$$[\frac{2n\hat{\beta}}{\chi^2_{\alpha/2,2n}}, \frac{2n\hat{\beta}}{\chi^2_{1-\alpha/2,2n}}] = [\frac{2 \cdot 5 \cdot 1274}{129.561}, \ \ \frac{2 \cdot 5 \cdot 1274}{74.222}] = \underline{\underline{[983,\ \ 1716]}}.$$

f) I point d) when we only have 5 observations there is a large difference between the exact interval and the approximate Wald interval. In point e) when we have 50 observations there is far less difference between the intervals. The approximate interval is much better in this case. Also notice that the intervals are much wider when we have only 5 observation than when we have 50 observations, reflecting that we have less information in the first case.

Exercise 1:

a) Since $\hat{\mu} = \frac{6}{7}\bar{X} + \frac{1}{7}\bar{Y}$ is a linear combination of independent normally distributed variables $\hat{\mu}$ will also be normally distributed. In 2 a) in exercise set 4 we found that $\mathrm{E}(\hat{\mu}) = \mu$ and $\mathrm{Var}(\hat{\mu}) = 0.000857$. We thus get:

$$Z = \frac{\hat{\mu} - \mathrm{E}(\hat{\mu})}{\sqrt{\mathrm{Var}(\hat{\mu})}} = \frac{\hat{\mu} - \mu}{\sqrt{0.000857}} \sim N(0,1)$$

$$\begin{aligned}
P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) &= 1 - \alpha \\
P(-z_{\alpha/2} \leq \frac{\hat{\mu} - \mu}{\sqrt{0.000857}} \leq z_{\alpha/2}) &= 1 - \alpha \\
P(-z_{\alpha/2}\sqrt{0.000857} \leq \hat{\mu} - \mu \leq z_{\alpha/2}\sqrt{0.000857}) &= 1 - \alpha \\
P(\hat{\mu} - z_{\alpha/2}\sqrt{0.000857} \leq \mu \leq \hat{\mu} + z_{\alpha/2}\sqrt{0.000857}) &= 1 - \alpha
\end{aligned}$$

Inserted $\alpha = 0.05 \ \ \Rightarrow \ \ z_{\alpha/2} = z_{0.025} = 1.96$ and $\hat{\mu} = \frac{6}{7} \cdot 6.12 + \frac{1}{7} \cdot 6.05 = 6.11$ this gives that a 95% confidence interval for μ is given by:

$$[6.11 - 1.96\sqrt{0.000857},\ 6.11 + 1.96\sqrt{0.000857}] = \underline{\underline{[6.05,\ 6.17]}}$$

Exercise 2:

a) Since $\hat{\mu}$ is a linear combination of the two independent normally distributed variables X_1 and X_2 it follows that $\hat{\mu}$ also is normally distributed. Using the expectation and variance calculated in point 3c) in exercise set 4 we get that

$$\frac{\hat{\mu} - \mathrm{E}(\hat{\mu})}{\sqrt{\mathrm{Var}(\hat{\mu})}} = \frac{\hat{\mu} - \mu}{\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}}} \sim N(0,1)$$

and a confidence interval is then found by:

$$\begin{aligned} P(-z_{\alpha/2} \leq \frac{\hat{\mu} - \mu}{\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}}} \leq z_{\alpha/2}) &= 1-\alpha \\ P(-z_{\alpha/2}\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}} \leq \hat{\mu} - \mu \leq z_{\alpha/2}\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}}) &= 1-\alpha \\ P(\hat{\mu} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}} \leq \mu \leq \hat{\mu} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}}) &= 1-\alpha \end{aligned}$$

With $\alpha = 0.05$ we get that $z_{\alpha/2} = z_{0.025} = 1.96$ and a 95% confidence interval is given by

$$\underline{\underline{\left[\hat{\mu} - 1.96\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}},\ \hat{\mu} + 1.96\sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}}\right]}}$$

Inserted observed numbers we get $\hat{\mu} = \frac{0.2^2 \cdot 4.41 + 0.5^2 \cdot 4.19}{0.5^2+0.2^2} = 4.220$ and $\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} = \frac{0.5^2 \cdot 0.2^2}{0.5^2+0.2^2} = 0.0345$ which gives the confidence interval:

$$[4.220 - 1.96 \cdot \sqrt{0.0345},\ 4.220 + 1.96 \cdot \sqrt{0.0345}] = \underline{\underline{[3.86, 4.58]}}$$

Exercise 3:

a)

$$\begin{aligned} P(X > 4) &= \int_4^\infty \frac{2x}{3} e^{-x^2/3} = [-e^{-x^2/3}]_4^\infty = e^{-4^2/3} = \underline{\underline{0.0048}} \\ p &= P(X_1 > 4 \cap X_2 > 4) \overset{\text{indep.}}{=} P(X_1 > 4)P(X_2 > 4) = e^{-4^2/3}e^{-4^2/3} = \underline{\underline{0.000023}} \end{aligned}$$

Generally

$$p = P(X_2 > 4 \cap X_1 > 4) = P(X_2 > 4 | X_1 > 4)P(X_1 > 4)$$

where $P(X_2 > 4|X_1 > 4) > P(X_2 > 4)$ if X_1 and X_2 are positively correlated, i.e. $\underline{\underline{p \text{ will increase}}}$ if we have a positive correlation between the wave heights.

b)

$$\begin{aligned}
L(\theta) &= \prod_{i=1}^{n}(\frac{2x_i}{\theta}e^{-x_i^2/\theta}) = \frac{2^n}{\theta^n}(\prod_{i=1}^{n}x_i)e^{-\sum_{i=1}^{n}x_i^2/\theta} \\
l(\theta) &= \ln L(\theta) = \ln(2^n) - \ln(\theta^n) + \ln(\prod_{i=1}^{n}x_i) + \ln e^{-\sum_{i=1}^{n}x_i^2/\theta} \\
&= n\ln(2) - n\ln(\theta) + \ln(\prod_{i=1}^{n}x_i) - \frac{\sum_{i=1}^{n}x_i^2}{\theta} \\
\frac{\partial l(\theta)}{\partial\theta} &= -\frac{n}{\theta} + \frac{\sum_{i=1}^{n}x_i^2}{\theta^2} = 0 \qquad \Rightarrow \hat{\theta} = \underline{\underline{\frac{1}{n}\sum_{i=1}^{n}X_i^2}}
\end{aligned}$$

c) Comparing the pdf of the Rayleigh distribution to the pdf of the Weibull distribution we see that the Rayleigh distribution is a special case of the Weibull distribution with parameters $\alpha = 1/\theta$ and $\beta = 2$. From the transformation result (collection of formulas) that if X is having a Weibull distribution with parameters α and β then $Y = 2\alpha X^\beta$ is having a χ_2^2-distribution we here get that $Z = 2(1/\theta)X^2 = 2X^2/\theta$ has a $\underline{\underline{\chi_2^2\text{-distribution}}}$.

Notice that $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}X_i^2 = \frac{\theta}{2n}\sum_{i=1}^{n}\frac{2X_i^2}{\theta} = \frac{\theta}{2n}\sum_{i=1}^{n}Z_i$ were we have from the above that the Z_is are χ_2^2-distributed. Further we have from the results on the expectation and the variance in the χ^2-distribution that $\mathrm{E}(Z_i) = 2$ and $\mathrm{Var}(Z_i) = 2 \cdot 2 = 4$ when Z_i is χ_2^2-distributed. We then get:

$$\begin{aligned}
\mathrm{E}(\hat{\theta}) &= \mathrm{E}(\frac{\theta}{2n}\sum_{i=1}^{n}Z_i) = \frac{\theta}{2n}\sum_{i=1}^{n}\mathrm{E}(Z_i) = \frac{\theta}{2n}\sum_{i=1}^{n}2 = \frac{\theta}{2n}2n = \underline{\underline{\theta}} \\
\mathrm{Var}(\hat{\theta}) &= \mathrm{Var}(\frac{\theta}{2n}\sum_{i=1}^{n}Z_i) = \frac{\theta^2}{(2n)^2}\sum_{i=1}^{n}\mathrm{Var}(Z_i) = \frac{\theta^2}{(2n)^2}\sum_{i=1}^{n}4 = \frac{\theta^2}{(2n)^2}4n = \underline{\underline{\frac{\theta^2}{n}}}
\end{aligned}$$

d) We see that $\frac{2n\hat{\theta}}{\theta} = \frac{2n}{\theta}\frac{1}{n}\sum_{i=1}^{n}X_i^2 = \sum_{i=1}^{n}\frac{2}{\theta}X_i^2 = \sum_{i=1}^{n}Z_i$ where $Z_i \sim \chi_2^2$. Since a sum of n χ_2^2-distributed variables is having a χ_{2n}^2-distribution (collection of formulas) we have that $\frac{2n\hat{\theta}}{\theta} = \sum_{i=1}^{n}Z_i \sim \underline{\underline{\chi_{2n}^2}}$.

Confidence interval:

$$\begin{aligned}
P(\chi^2_{1-\alpha/2,2n} \le \frac{2n\hat{\theta}}{\theta} \le \chi^2_{\alpha/2,2n}) &= 1-\alpha \\
P(\frac{\chi^2_{1-\alpha/2,2n}}{2n\hat{\theta}} \le \frac{1}{\theta} \le \frac{\chi^2_{\alpha/2,2n}}{2n\hat{\theta}}) &= 1-\alpha \\
P(\frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}} \ge \theta \ge \frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}) &= 1-\alpha \\
P(\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}} \le \theta \le \frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}) &= 1-\alpha
\end{aligned}$$

I.e. a $(1-\alpha)100\%$ confidence interval for θ is give by:

$$\underline{\underline{\left[\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}},\ \frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}\right]}}$$

With $n=50$, $\hat{\theta}=1.047$ and $\alpha=0.05$ which gives $\chi^2_{1-\alpha/2,2n}=\chi^2_{0.975,100}=74.222$ and $\chi^2_{\alpha/2,2n}=\chi^2_{0.025,100}=129.561$ we get:

$$\left[\frac{2\cdot 50\cdot 1.047}{129.561},\ \frac{2\cdot 50\cdot 1.047}{74.222}\right]=\underline{\underline{[0.81,\ \ 1.41]}}$$

e) Starting from the interval in d) we get:

$$\begin{aligned}
P\left(\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}\leq\theta\leq\frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}\right) &= 1-\alpha\\
P\left(\sqrt{\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}\pi}\leq\sqrt{\theta\pi}\leq\sqrt{\frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}\pi}\right) &= 1-\alpha\\
P\left(0.5\sqrt{\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}\pi}\leq 0.5\sqrt{\theta\pi}\leq 0.5\sqrt{\frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}\pi}\right) &= 1-\alpha\\
P\left(0.5\sqrt{\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}\pi}\leq\mu\leq 0.5\sqrt{\frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}\pi}\right) &= 1-\alpha
\end{aligned}$$

I.e. a $(1-\alpha)100\%$ confidence interval for μ is give by:

$$\underline{\underline{\left[0.5\sqrt{\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}\pi},\ 0.5\sqrt{\frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}\pi}\right]}}$$

With $n=50$, $\hat{\theta}=1.047$, $\chi^2_{1-\alpha/2,2n}=\chi^2_{0.975,100}=74.222$ and $\chi^2_{\alpha/2,2n}=\chi^2_{0.025,100}=129.561$ we get:

$$\left[0.5\sqrt{\frac{2\cdot 50\cdot 1.047}{129.561}\pi},\ 0.5\sqrt{\frac{2\cdot 50\cdot 1.047}{74.222}\pi}\right]=\underline{\underline{[0.80,\ \ 1.05]}}$$

f) We start by finding the second derivative of the log-likelihood at $\hat{\theta}$:

$$\begin{aligned}
\frac{\partial l(\theta)}{\partial\theta} &= -\frac{n}{\theta}+\frac{\sum_{i=1}^n x_i^2}{\theta^2}\\
J(\theta) &= \frac{\partial^2 l(\theta)}{\partial\theta^2}=\frac{n}{\theta^2}-2\frac{\sum_{i=1}^n x_i^2}{\theta^3}\\
J(\hat{\theta}) &= \frac{n}{\hat{\theta}^2}-2n\frac{\sum_{i=1}^n x_i^2}{n}\frac{1}{\hat{\theta}^3}=\frac{n}{\hat{\theta}^2}-2n\frac{1}{\hat{\theta}^2}=-\frac{n}{\hat{\theta}^2}
\end{aligned}$$

The Wald confidence interval then becomes:

$$[\hat{\theta} - z_{\alpha/2}\sqrt{-1/J(\hat{\theta})}, \hat{\theta} + z_{\alpha/2}\sqrt{-1/J(\hat{\theta})}] = [\hat{\theta} - z_{\alpha/2}\sqrt{\hat{\theta}^2/n}, \hat{\theta} + z_{\alpha/2}\sqrt{\hat{\theta}^2/n}]$$

and inserting $z_{0.025} = 1.96$, $n = 50$ and $\hat{\theta} = 1.047$ we get:

$$[1.047 - 1.96\sqrt{1.047^2/50}, \;\; 1.047 + 1.96\sqrt{1.047^2/50}] = \underline{\underline{[0.76, \;\; 1.34]}}$$

This is not very different from the exact interval in d), with $n = 50$ observations we have enough data for this approximate interval to be quite good.

g)

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(\max(X_1, \ldots X_m) \leq y) \\ &= P(X_1 \leq y \cap \cdots \cap X_m \leq y) \overset{indep.}{=} P(X_1 \leq y) \cdots P(X_m \leq y) = F_X(y)^m \end{aligned}$$

Since

$$F_X(y) = \int_{-\infty}^{y} f_X(x)dx = \int_0^y \frac{2x}{\theta} e^{-x^2/\theta} dx = 1 - e^{-y^2/\theta}$$

we get

$$F_Y(y) = \underline{\underline{[1 - e^{-y^2/\theta}]^m}}$$

Since $P(Y > y_c) = 1/100$ then $F_Y(y_c) = P(Y \leq y_c) = 1 - 1/100$:

$$\begin{aligned} F_Y(y_c) = \left[1 - e^{-y_c^2/\theta}\right]^m &= 1 - \frac{1}{100} \\ 1 - e^{-y_c^2/\theta} &= \left[1 - \frac{1}{100}\right]^{1/m} \\ 1 - \left[1 - \frac{1}{100}\right]^{1/m} &= e^{-y_c^2/\theta} \\ \ln\left(1 - \left[1 - \frac{1}{100}\right]^{1/m}\right) &= -y_c^2/\theta \\ y_c &= \underline{\underline{\sqrt{-\theta \ln\left(1 - \left[1 - \frac{1}{100}\right]^{1/m}\right)}}} \end{aligned}$$

h) If $y_c < 4$ the probability of getting a wave higher than 4 meters is less than 0.01. The operation starts if the measurements give basis to claim that $y_c < 4$. In the previous point we found the relation between y_c and the parameter θ. We can use this relation and the confidence interval for θ from point d) to calculate a confidence interval for y_c. To simplify the notation we insert $m = 900$ and get the relationship written as

$$y_c = \sqrt{-\theta \ln\left(1 - \left[1 - \frac{1}{100}\right]^{1/900}\right)} = \sqrt{-\theta \cdot (-11.4)} = \sqrt{11.4 \cdot \theta}$$

The confidence interval then becomes:

$$\begin{aligned} P\left(\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}} \leq \theta \leq \frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}\right) &= 1-\alpha \\ P\left(\sqrt{11.4 \cdot \frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}} \leq \sqrt{11.4 \cdot \theta} \leq \sqrt{11.4 \cdot \frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}}\right) &= 1-\alpha \\ P\left(\sqrt{11.4 \cdot \frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}} \leq y_c \leq \sqrt{11.4 \cdot \frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}}\right) &= 1-\alpha \end{aligned}$$

Inserting $n = 50$, $\hat{\theta} = 1.047$ and $\alpha = 0.10$ which gives $\chi^2_{1-\alpha/2,2n} = \chi^2_{0.95,100} = 77.929$ and $\chi^2_{\alpha/2,2n} = \chi^2_{0.05,100} = 124.342$ we get:

$$\left[\sqrt{11.4\frac{2 \cdot 50 \cdot 1.047}{124.342}}, \sqrt{11.4\frac{2 \cdot 50 \cdot 1.047}{77.929}}\right] = \underline{\underline{[3.10, \quad 3.91]}}$$

We see that the upper limit of this interval is smaller than 4 and we can then claim that $y_c < 4$. The operation can begin!