Université d'Ottawa · University of Ottawa
Faculté des sciences
Mathématiques et de statistique
Faculty of Science
Mathematics and Statistics

Solution to Assignment 4, Fall 2014
Probability and Statistics for Engineers.

585, av. King-Edward
Ottawa (Ontario) K1N 6N5 Canada
585 King Edward Avenue
Ottawa, Ontario K1N 6N5 Canada
(613) 562-5864 • Téléc./Fax (613) 562-5776
Courriel/Email: uomaths@science.uottawa.ca

1. Textbook problems.

3.94. For $1 \leq i \leq 100$, let X_i be 1 if the i'th switch fails within a year and 0 otherwise. These are Bernoulli random variables, and

$$\mathbb{P}[X_i = 1] = 1 - e^{-\frac{1}{2}} \approx 0.393.$$

Let $X = \sum_{i=1}^{100} X_i$. We need to calculate $\mathbb{P}[X \leq 30]$. Using the normal approximation to the binomial, we have

$$\mathbb{P}[X \leq 30] = \mathbb{P}[\frac{X - 39.3}{\sqrt{100(0.393)(1-0.393)}} \leq \frac{30 - 39.3}{\sqrt{100(0.393)(1-0.393)}}] \approx \mathbb{P}[Z \leq -1.904] \approx 0.028.$$

3.110. Let X have negative binomial distribution with parameters $k = 2$ and $p = 0.25$. This problem asks us to calculate $\mathbb{P}[X \leq 6]$. We calculate :

$$\mathbb{P}[X \leq 6] = \sum_{i=2}^{6} \binom{i-1}{1} (0.25)^2 (0.75)^{i-2} \approx 0.466.$$

3.136.

(a) Let X be the time until the next breakdown. By the 'memoryless property',

$$\mathbb{P}[X \leq 21] = 1 - e^{-\frac{21}{15}} \approx 0.753.$$

(b) Doing a similar calculation,

$$\mathbb{P}[X \geq 30] = e^{-\frac{30}{15}} \approx 0.135.$$

4.10. For both parts, we consider a sample $X = (x_1, x_2, \ldots, x_n)$.

(a) Let $Y = (y_1, \ldots, y_n) = (x_1 + c, \ldots, x_n + c)$ be our shifted sample. We note that

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}(x_i + x) = \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=1}^{n} c = \overline{X} + c.$$

Let S_Y^2, S_X^2 be the sample variances for Y and X respectively. Then

$$S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{Y})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i + c - \overline{X} - c)^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2 = S_X^2.$$

This completes the proof.

(b) Let $Y = (y_1, \ldots, y_n) = (cx_1, \ldots, cx_n)$ be the shifted sample for the second part. Then

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n} cx_i = \frac{c}{n}\sum_{i=1}^{n} x_i = c\overline{X}.$$

Let S_Y^2, S_X^2 be the sample variances for Y and X respectively. Then

$$S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{Y})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(cx_i - c\overline{X})^2 = \frac{c^2}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2 = c^2 S_X^2.$$

This completes the proof.

4.17.

(a) We calculate directly

$$\mu = (0.2)(4) + (0.4)(5) + (0.3)(6) + (0.1)(7) = 5.3$$

and

$$\sigma^2 = (0.2)(4-5.3)^2 + (0.4)(5-5.3)^2 + (0.3)(6-5.3)^2 + (0.1)(7-5.3)^2 = 0.81.$$

(b) We use the formulae from section 4.4 :

$$\mu_{\overline{X}} = \mu = 5.3$$

and

$$\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} = \frac{0.81}{36} = 0.0225.$$

(c) By the central limit theorem,

$$\mathbb{P}[\overline{X} < 5.5] = \mathbb{P}[\frac{\overline{X} - 5.3}{\sqrt{0.0225}} < \frac{5.5 - 5.3}{\sqrt{0.0225}}] \approx \mathbb{P}[Z \leq 1.333] \approx 0.909.$$

2.

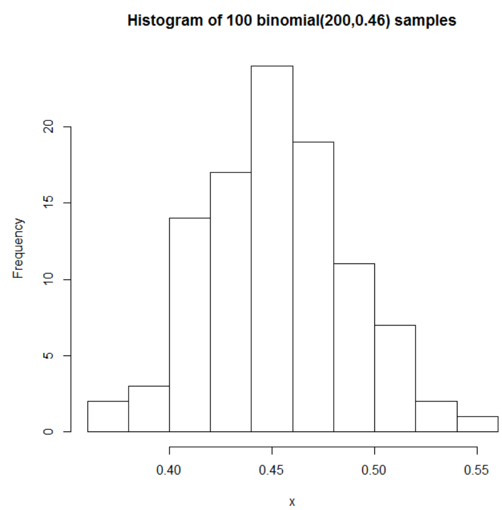
(a) We use the R code : x = *rnorm(60,10,5)*, then *summary(x)*, *mean(x)*, and *sqrt((60 \ 59) ∗ var(x))*. For me, the results were :

Min = -1.931, 1st Quartile = 5.626, Median = 11.090, Mean = 10.030, 3rd Quartile = 13.810, Max = 24.040.
$\mu \approx 10.03$
$\sigma \approx 5.39$.

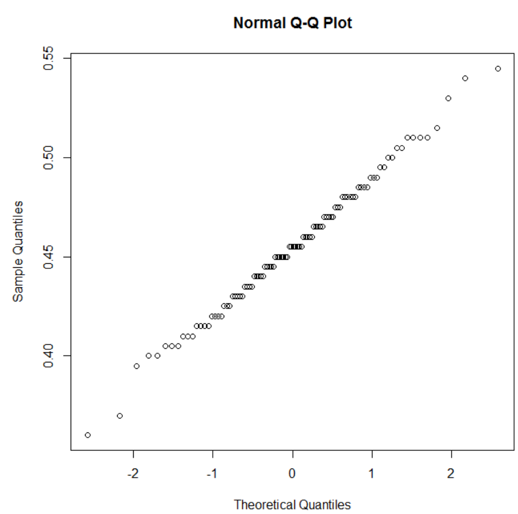
NOTE : It is ok to not use the correction factor $\frac{60}{59}$ in this question, or to use a slightly different correction such as $\frac{60}{61}$ or even $\frac{60}{62}$. As discussed in class, these estimator have some (different) desirable properties, and all are quite close for reasonable amounts of data.

(b) I omit the R code in this pdf (see other notes for similar computations). When I ran the program twice, I found that 2 of the 50 (i.e. four percent) were not in the confidence interval the first time and 0 of the 50 (i.e. 0 percent) were not in the confidence interval the second time.

3. First, the histogram :



Then the qq plot :



These are meant to illustrate the Central Limit Theorem.