STA500 Introduction to Probability and Statistics 2, autumn 2018.

Solution exercise set 4

Exercises from the book:

8.13

The easiest way to do this is to just use the function for calculating sample standard deviations on your calculator. Then you just punch in the data on the calculator and get the sample standard deviation calculated.

Alternatively you can approach the problem by first simplify the formula for s^2, for instance to (see also theorem 8.1 in the textbook):

$$\begin{aligned} s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i^2-2\bar{x}x_i+\bar{x}^2) = \frac{1}{n-1}(\sum_{i=1}^{n}x_i^2-2\bar{x}\sum_{i=1}^{n}x_i+n\bar{x}^2) \\ &= \frac{1}{n-1}(\sum_{i=1}^{n}x_i^2-2\bar{x}n\bar{x}+n\bar{x}^2) = \frac{1}{n-1}(\sum_{i=1}^{n}x_i^2-n\bar{x}^2) \end{aligned}$$

For the given data we get $\bar{x} = \sum_{i=1}^{20} x_i/20 = 53.3/20 = 2.665$ and $\sum_{i=1}^{20} x_i^2 = 148.55$ which gives

$$s^2 = \frac{1}{20-1}(148.55 - 20\cdot 2.665^2) = 0.342$$

The *sample* (or *estimated* or *empirical*) standard deviation (the book wrongly writes just "the standard deviation" in the exercise text) then becomes: $s = \sqrt{0.342} = \underline{\underline{0.585}}$.

8.14

a) If we add a constant c to all the data $x_1, \ldots, x_n$, the data becomes $x_1 + c, \ldots, x_n + c$, and the average becomes $\bar{x} + c$. We then get:

$$S_{X+c}^2 = \frac{1}{n-1}\sum_{i=1}^{n}((x_i+c)-(\bar{x}+c))^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \underline{\underline{S_X^2}}$$

b) If we multiply all data $x_1, \ldots, x_n$ by a constant c, the data becomes $cx_1, \ldots, cx_n$, and the average becomes $c\bar{x}$. We then get:

$$S_{cX}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(cx_i-c\bar{x})^2 = \frac{c^2}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \underline{\underline{c^2S_X^2}}$$

8.21

Since $\sigma_{\bar{X}} = \mathrm{SD}(\bar{X}) = \sqrt{\mathrm{Var}(\bar{X})} = \sqrt{\sigma^2/n} = \sigma/\sqrt{n}$ and $\mu_{\bar{X}} = \mathrm{E}(\bar{X}) = \mu$ with $n = 40$, $\sigma = 15$ and $\mu = 240$ (which holds if the machine is correctly adjusted!) we get that the machine is correctly adjusted if 40 is in the interval

$$[\mu_{\bar{X}} - 2\sigma_{\bar{X}}, \mu_{\bar{X}} + 2\sigma_{\bar{X}}] = [240 - 2 \cdot 15/\sqrt{40}, 240 + 2 \cdot 15/\sqrt{40}] = [235.3, 244.7]$$

Since 236 is in this interval they made the $\underline{\underline{\text{correct decision}}}$.

Notice that from the rule of thumb saying that in many cases around 95% of the measurements will fall in an interval plus/minus two standard deviations from the expectation, we get in the present example that if the machine is correctly adjusted there is a probability of around 95% of getting a value in the calculated interval.

8.26

Let X be the amount of time spent by a random customer. Notice that the distribution of X has not be specified! However, we know that X is having expectation $\mu = 3.2$ and standard deviation $\sigma = 1.6$. Further let $\bar{X} = \frac{1}{64}\sum_{i=1}^{64} X_i$ be the average time spent by 64 randomly selected customers. The central limit theorem (CLT) now gives that:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - 3.2}{1.6/8} = \frac{\bar{X} - 3.2}{0.2} \approx N(0,1)$$

i.e. Z is having an approximate normal distribution.

a)

$$\begin{aligned} P(\bar{X} \le 2.7) &= P(\frac{\bar{X} - 3.2}{0.2} \le \frac{2.7 - 3.2}{0.2}) \\ &= P(Z \le -2.5) \\ &\overset{CLT}{\approx} \underline{\underline{0.0062}} \end{aligned}$$

b)

$$\begin{aligned} P(\bar{X} > 3.5) &= 1 - P(Z \le \frac{3.5 - 3.2}{0.2}) \\ &\overset{CLT}{\approx} 1 - 0.9332 \\ &= \underline{\underline{0.0668}} \end{aligned}$$

c)

$$\begin{aligned} P(3.2 \le \bar{X} \le 3.4) &= P(\bar{X} < 3.4) - P(\bar{X} < 3.2) \\ &= P(Z < \frac{3.4 - 3.2}{0.2}) - P(Z < \frac{3.2 - 3.2}{0.2}) \\ &\overset{CLT}{\approx} 0.8413 - 0.5 \\ &= \underline{\underline{0.3413}} \end{aligned}$$

9.81

The exercise text is maybe a bit unclear in this exercise. When it is said that $x_1, \ldots, x_n$ is a Bernoulli-process (or the outcome of a binomial trial) with parameter p, this means that $x_1, \ldots, x_n$ is the outcome of a 0-1 variable $X_1, \ldots, X_n$ with distribution

$$f(x) = p^x(1-p)^{(1-x)} \quad , x = 0, 1$$

(i.e. $P(X=1) = p$ and $P(X=0) = 1-p$). We then get

$$\begin{aligned}
L(p; x_1, \ldots, x_n) &\overset{\text{indep.}}{=} \prod_{i=1}^{n} f(x_i, p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&= p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i} \\
l(p; x_i, \ldots, x_n) &= \ln L(p; x_i, \ldots, x_n) = \sum_{i=1}^{n} x_i \ln p + \left(n - \sum_{i=1}^{n} x_i\right) \ln(1-p) \\
\frac{\partial l(p; x_1, \ldots, x_n)}{\partial p} &= \frac{1}{p}\sum_{i=1}^{n} x_i - \frac{1}{1-p}\left(n - \sum_{i=1}^{n} x_i\right) = 0 \\
&\Rightarrow \quad (1-p)\sum_{i=1}^{n} x_i - p\left(n - \sum_{i=1}^{n} x_i\right) = 0 \\
& \sum_{i=1}^{n} x_i - p\sum_{i=1}^{n} x_i - pn + p\sum_{i=1}^{n} x_i = 0 \\
&\Rightarrow \quad \underline{\underline{\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i}}
\end{aligned}$$

Checking that we have found a maximum:

$$\frac{\partial^2 l(p; x_1, \ldots, x_n)}{\partial p^2} = -\frac{1}{p^2}\sum_{i=1}^{n} x_i - \frac{1}{(1-p)^2}\left(n - \sum_{i=1}^{n} x_i\right) < 0 \qquad \text{i.e. maximum!}$$

We can also check if the estimator is unbiased. Notice first that $\mathrm{E}(X) = \sum_x x f(x) = 0(1-p) + 1p = p$. Then:

$$E[\hat{p}] = E[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} EX_i = \frac{1}{n}\sum_{i=1}^{n} p = \underline{p}$$

We see that $E[\hat{p}] = p$, the estimator is unbiased.

9.84/9.82

a)

$$\begin{aligned}
L(\alpha, \beta; x_1, ..x_n) &= \prod_{i=1}^{n} f(x_i; \alpha, \beta) = \prod_{i=1}^{n} \alpha\beta x_i^{\beta-1} e^{-\alpha x_i^\beta} \\
&= (\alpha\beta)^n (\prod_{i=1}^{n} x_i^{\beta-1}) e^{-\sum_{i=1}^{n} \alpha x_i^\beta} = \underline{\underline{(\alpha\beta)^n e^{-\alpha \sum_{i=1}^{n} x_i^\beta} \prod_{i=1}^{n} x_i^{\beta-1}}}
\end{aligned}$$

b)

$$\begin{aligned} l(\alpha,\beta;x_1,..x_n) &= \ln L(\alpha,\beta;x_1,..x_n) = n\ln(\alpha)+n\ln(\beta)-\alpha\sum_{i=1}^{n}x_i^{\beta}+\sum_{i=1}^{n}\ln(x_i^{\beta-1}) \\ &= n\ln(\alpha)+n\ln(\beta)-\alpha\sum_{i=1}^{n}x_i^{\beta}+(\beta-1)\sum_{i=1}^{n}\ln(x_i) \end{aligned}$$

Since we here shall optimize over two parameters, α and β we must solve the equations $\frac{\partial l(\alpha,\beta)}{\partial\alpha}=0$ and $\frac{\partial l(\alpha,\beta)}{\partial\beta}=0$ simultaneously, i.e. solve the system of equations:

$$\begin{aligned} \frac{\partial l(\alpha,\beta)}{\partial\alpha} &= \frac{n}{\alpha}-\sum_{i=1}^{n}x_i^{\beta}=0 \\ \frac{\partial l(\alpha,\beta)}{\partial\beta} &= \frac{n}{\beta}-\alpha\sum_{i=1}^{n}x_i^{\beta}\ln(x_i)+\sum_{i=1}^{n}\ln(x_i)=0 \end{aligned}$$

The first equation gives a simple expression for α ($\alpha = n/\sum_{i=1}^{n}x_i^{\beta}$) which can be inserted in the second equation, and we can then try to solve this equation with respect to β. However, this equation can not be solved analytically, numerical methods must be applied to find a solution.

Exercise 1:

We first calculate the cumulative distribution function of X:

$$F_X(x)=\int_0^x \lambda e^{-\lambda t}dt = 1-e^{-\lambda x} \qquad \text{for} \qquad x>0$$

To find the probability density of $V=\max(X_1,X_2)$ we first calculate the cumulative distribution function:

$$\begin{aligned} F_V(v)=P(\max(X_1,X_2)\le v) &= P(X_1\le v\cap X_2\le v) \overset{indep.}{=} P(X_1\le v)P(X_2\le v) \\ &= F_X(v)^2=(1-e^{-\lambda v})^2=1-2e^{-\lambda v}+e^{-2\lambda v} \qquad \text{for} \qquad v>0 \end{aligned}$$

I.e. the pdf of V becomes:

$$f_V(v)=F'(v)=\underline{\underline{2\lambda e^{-\lambda v}-2\lambda e^{-2\lambda v}}} \qquad \text{for} \qquad v>0$$

For the exponential distribution we have that $\mathrm{E}(X)=\int_0^\infty x\lambda e^{-\lambda x}dx=\frac{1}{\lambda}$, and we can use this result also in the calculation of $E(V)$:

$$\begin{aligned} \mathrm{E}(V)=\int_{-\infty}^{\infty}vf(v)dv=\int_0^\infty v(2\lambda e^{-\lambda v}-2\lambda e^{-2\lambda v})dv &= 2\int_0^\infty v\lambda e^{-\lambda v}dv-\int_0^\infty v2\lambda e^{-2\lambda v}dv \\ &= 2\frac{1}{\lambda}-\frac{1}{2\lambda}=\underline{\underline{\frac{3}{2\lambda}}} \end{aligned}$$

Thus since $\mathrm{E}(X)=\frac{1}{\lambda}$ we have that $\underline{\underline{\mathrm{E}(X)<\mathrm{E}(V)<2\mathrm{E}(X)}}$ which is as expected since V is the largest of to X-values. Since $V=\max(X_1,X_2)$ we will expect that $\mathrm{E}(V)>\mathrm{E}(X)$ and that $\mathrm{E}(V)<\mathrm{E}(X_1+X_2)=2\mathrm{E}(X)$.

Exercise 2:

a)

$$\begin{aligned}
\mathrm{E}(\hat{\mu}_1) &= \frac{1}{25}[\sum_{i=1}^{10}\mathrm{E}(X_i)+\sum_{i=1}^{15}\mathrm{E}(Y_i)] = \frac{1}{25}[\sum_{i=1}^{10}\mu+\sum_{i=1}^{15}\mu] = \frac{1}{25}[10\mu+15\mu] = \underline{\underline{\mu}} \\
\mathrm{E}(\hat{\mu}_2) &= \frac{6}{7}\mathrm{E}(\bar{X})+\frac{1}{7}\mathrm{E}(\bar{Y}) = \frac{6}{7}\mu+\frac{1}{7}\mu = \underline{\underline{\mu}}
\end{aligned}$$

I.e. both estimators are unbiased.

$$\begin{aligned}
\mathrm{Var}(\hat{\mu}_1) &= \mathrm{Var}(\frac{1}{25}[\sum_{i=1}^{10}X_i+\sum_{i=1}^{15}Y_i]) = \frac{1}{25^2}\mathrm{Var}[\sum_{i=1}^{10}X_i+\sum_{i=1}^{15}Y_i] \\
&\overset{\text{indep.}}{=} \frac{1}{25^2}[\sum_{i=1}^{10}\mathrm{Var}(X_i)+\sum_{i=1}^{15}\mathrm{Var}(Y_i)] = \frac{1}{25^2}[10\cdot 0.01+15\cdot 0.09] = \underline{\underline{0.00232}} \\
\mathrm{Var}(\hat{\mu}_2) &= \mathrm{Var}(\frac{6}{7}\bar{X}+\frac{1}{7}\bar{Y}) \overset{\text{indep.}}{=} (\frac{6}{7})^2\mathrm{Var}(\bar{X})+(\frac{1}{7})^2\mathrm{Var}(\bar{Y}) \\
&= (\frac{6}{7})^2\frac{0.01}{10}+(\frac{1}{7})^2\frac{0.09}{15} = \underline{\underline{0.000857}}
\end{aligned}$$

Since both estimators are unbiased and $\mathrm{Var}(\hat{\mu}_1) > \mathrm{Var}(\hat{\mu}_2)$, $\underline{\underline{\hat{\mu}_2 \text{ is the best estimator}}}$.

b) Since we want to have an unbiased estimator, we must have:

$$\mathrm{E}(\hat{\mu}) = a\mathrm{E}(\bar{X})+b\mathrm{E}(\bar{Y}) = a\mu+b\mu = \mu$$

I.e. we must have $a+b=1$. Further we want to have an estimator with as small variance as possible. The variance of the estimator becomes:

$$\begin{aligned}
\mathrm{Var}(\hat{\mu}) &\overset{\text{indep.}}{=} a^2\mathrm{Var}(\bar{X})+b^2\mathrm{Var}(\bar{Y}) = a^2\frac{0.01}{10}+b^2\frac{0.09}{15} \\
&= 0.001a^2+0.006b^2
\end{aligned}$$

I.e. to find the unbiased estimator with the smallest possible variance, we must find the values a and b where $a+b=1$ and where at the same time $0.001a^2+0.006b^2$ is as small as possible. Substituting $b=1-a$ into the last expression we get the following expression to minimize:

$$\begin{aligned}
V(a) &= 0.001a^2+0.006(1-a)^2 = 0.007a^2-0.012a+0.006 \\
V'(a) &= 0.014a-0.012 = 0 \\
&\Rightarrow a = \frac{0.012}{0.014} = \underline{\underline{\frac{6}{7}}} \\
&\Rightarrow b = 1-\frac{6}{7} = \underline{\underline{\frac{1}{7}}}
\end{aligned}$$

Hence we see that $\hat{\mu}_2$ from b) is the best unbiased estimator in the present case!

Exercise 3:

Let X_1 denote the result of a measurement using method 1 and let X_2 denote the result of a measurement using method 2.

a)

$$
\begin{aligned}
P(X_1 < 4.0) &= P(Z < \frac{4.0-4.3}{0.5}) = P(Z < -0.60) = \underline{\underline{0.2743}} \\
P(X_2 > \mu + \sigma_2) &= 1 - P(X_2 \le \mu + \sigma_2) = 1 - P(Z \le \frac{\mu + \sigma_2 - \mu}{\sigma_2}) \\
&= 1 - P(Z \le 1) = 1 - 0.8413 = \underline{\underline{0.1587}}
\end{aligned}
$$

b)

$$
\begin{aligned}
L(\mu) &= \prod_{i=1}^{n} f(x_i;\mu) = f(x_1;\mu) f(x_2;\mu) \\
&= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(x_1-\mu)^2/(2\sigma_1^2)} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-(x_2-\mu)^2/(2\sigma_2^2)} \\
&= \frac{1}{2\pi\sigma_1\sigma_2} e^{-(x_1-\mu)^2/(2\sigma_1^2)-(x_2-\mu)^2/(2\sigma_2^2)} \\
l(\mu) &= \ln L(\mu) = \ln(1) - \ln(2\pi\sigma_1\sigma_2) - \frac{(x_1-\mu)^2}{2\sigma_1^2} - \frac{(x_2-\mu)^2}{2\sigma_2^2}
\end{aligned}
$$

Take the derivative with respect to μ and set equal to zero:

$$
\begin{aligned}
\frac{\partial l(\mu)}{\partial \mu} = -\frac{2(x_1-\mu)(-1)}{2\sigma_1^2} - \frac{2(x_2-\mu)(-1)}{2\sigma_2^2} &= 0 \\
\frac{x_1-\mu}{\sigma_1^2} + \frac{x_2-\mu}{\sigma_2^2} &= 0 \\
\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2} &= \frac{\mu}{\sigma_1^2} + \frac{\mu}{\sigma_2^2} \\
\sigma_2^2 x_1 + \sigma_1^2 x_2 &= \sigma_2^2\mu + \sigma_1^2\mu \\
\hat{\mu} &= \underline{\underline{\frac{\sigma_2^2 X_1 + \sigma_1^2 X_2}{\sigma_1^2 + \sigma_2^2}}}
\end{aligned}
$$

c)

$$
\mathrm{E}(\hat{\mu}) = \mathrm{E}\left(\frac{\sigma_2^2 X_1 + \sigma_1^2 X_2}{\sigma_1^2 + \sigma_2^2}\right) = \frac{\sigma_2^2 \mathrm{E}(X_1) + \sigma_1^2 \mathrm{E}(X_2)}{\sigma_1^2 + \sigma_2^2} = \frac{\sigma_2^2\mu + \sigma_1^2\mu}{\sigma_1^2 + \sigma_2^2} = \underline{\underline{\mu}}
$$

I.e. the estimator is unbiased.

$$
\begin{aligned}
\mathrm{Var}(\hat{\mu}) &= \mathrm{Var}\left(\frac{\sigma_2^2 X_1 + \sigma_1^2 X_2}{\sigma_1^2 + \sigma_2^2}\right) = \mathrm{Var}\left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} X_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} X_2\right) \\
&= \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)^2 \mathrm{Var}(X_1) + \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2 \mathrm{Var}(X_2) = \frac{\sigma_2^4\sigma_1^2 + \sigma_1^4\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)^2} \\
&= \frac{\sigma_1^2\sigma_2^2(\sigma_2^2 + \sigma_1^2)}{(\sigma_1^2 + \sigma_2^2)^2} = \underline{\underline{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}}}
\end{aligned}
$$