Solutions exercises chapter 7

1. **a**. $P(X > 90 \text{ and } Y > 90) \overset{\text{ind.}}{=} P(X > 90) \cdot P(Y > 90) = \left(1 - \Phi\left(\frac{90-80}{10}\right)\right)^2 \approx 2.52\%$

b. $X + Y$ is $N(80 + 80, 100 + 100)$-verdeeld, dus

$P(X + Y > 180) = P\left(Z > \frac{180-160}{\sqrt{200}}\right) \approx 1 - \Phi(1.41) = 7.93\%$

c. The event "$X > 90$ and $Y > 90$" is a subset of "$X + Y > 180$" (try e.g. $X = 120$ and $Y = 80$), that is why the first event has a smaller probability.

2. **a.** $f(x) = \begin{cases} 1 \ if & 0 \leq x \leq 1 \\ 0 & elsewhere \end{cases}$ $\quad$ $(sketch\ the\ density\ function\ and\ a\ point\ x\ between\ 0\ and\ 1)$,

So $F(x) = P(X \leq x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases}$

b. $F_M(m) = P(max(X_1, X_2, X_3) \leq m) \overset{\text{ind.}}{=} P(X_1 \leq m)P(X_2 \leq m)P(X_3 \leq m) = [F(m)]^3$

(note that $F(m)$ is the distribution function given in a.)

$f_M(m) = \frac{d}{dm}([F(m)]^3) = 3F(m)^2 \cdot f(m) = 3m^2 \cdot 1$, for $0 \leq m \leq 1$

(you might also substitute $F(m) = m$ in the first line and then differentiate with respect to m).

3. **a**. $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx = \int_0^z e^{-x} e^{-(z-x)} dx = \int_0^z e^{-z} dx = e^{-z} \cdot x|_{x=0}^{x=z}$

$= ze^{-z}$, for $z \geq 0$

(And $f_{X+Y}(z) = 0$, if $z < 0$)

b. $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx = \int_0^z e^{-x} \cdot 2e^{-2(z-x)} dx = \int_0^z 2e^{-2z} e^x dx$

$= 2e^{-2z} \cdot e^x|_{x=0}^{x=z} = 2e^{-z} - 2e^{-2z}$, *for* $z \geq 0$

(And $f_{X+Y}(z) = 0$, if $z < 0$)

c. $P(X > 1 \text{ en } Y < 1) \overset{\text{ind.}}{=} P(X > 1) \cdot P(Y < 1) = e^{-1} \cdot (1 - e^{-2 \cdot 1}) \approx 31.8\%$

4. $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx = \int_0^z \frac{1}{\sqrt{2\pi x}} e^{-\frac{1}{2}x} \cdot \frac{1}{\sqrt{2\pi x}} e^{-\frac{1}{2}(z-x)} \text{dx} = \frac{e^{-\frac{1}{2}z}}{2\pi} \int_0^z \frac{1}{\sqrt{x(z-x)}} dx$

The last integral equals π (given): $f_{X+Y}(z) = \frac{1}{2} e^{-\frac{1}{2}z}$, for $z > 0$.

A Chi-square distribution with 2 degrees of freedom is apparently the same as an Exp $\left(\frac{1}{2}\right)$*-distribution.*

5. **a**. Yes, the rule $E(X + Y) = E(X) + E(Y)$ is universal.

b. No, this is only the case if independence is valid, but that seems an unreasonable assumption in this case: the salaries of men and their women are usually related, because their level of education is related and e.g. because they might agree upon who is at home for the children more than the other.

6. We will apply the normal distribution of $X - Y \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2) = N(50 - 60, 100 + 36)$.

($\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2$ because of independence)

Then we have: $P(X > Y) = P(X - Y > 0) = 1 - \Phi\left(\frac{0-(-10)}{\sqrt{136}}\right) \approx 0.1965$,

so we expect about (rounded) 20 bottles of 100 will break.

7. **a**. We start by rewriting the probability: $P(3X \leq 2Y - 1) = P(3X - 2Y \leq -1)$.

Because of independence of X and Y $3X - 2Y$ (a "linear combination") is normally distributed as well, with parameters $\boldsymbol{\mu} = \boldsymbol{E(3X - 2Y)} = 3E(X) - 2E(Y) = 3 \cdot 4 - 2 \cdot 2 = 8$ and

$\boldsymbol{\sigma^2} = \boldsymbol{var(3X - 2Y)} \overset{\text{ind.}}{=} var(3X) + var(2Y) = 3^2 var(X) + 2^2 var(Y) = 9 \cdot 1 + 4 \cdot 4 = 25$

So $P(3X - 2Y \leq -1) = P\left(Z \leq \frac{-1-8}{\sqrt{25}}\right) = \Phi(-1.80) = 0.0359$

b. $\rho(X, 3X - 2Y) = \frac{cov(X,\ 3X-2Y)}{\sigma_X \sigma_{3X-2Y}} = \frac{3}{5}$

since: $cov(X, 3X - 2Y) = cov(X, 3X) - cov(X, 2Y) = 3 \cdot var(X) - 0 = 3$
(using: $cov(X, 2Y) = 2 \cdot cov(X, Y)$ and $cov(X, Y) \overset{\text{ind.}}{=} 0$)

and $\sigma_{3X-2Y} = \sqrt{var(3X - 2Y)} = \sqrt{25} = 5$.

8. **a**. Assuming that X and Y are independent we have:
$X + Y \sim N(75 + 65, 250 + 150)$, so $P(X + Y > 150) = 1 - \Phi\left(\frac{150-140}{\sqrt{400}}\right) = 1 - \Phi\left(\frac{1}{2}\right)$ (= 30.85%)
This is an exact computation, based on exact normal distributions. Of course the normal model for the weights might be an approximate model of the reality. (Within the model computations are exact)

b. According to CLT ($n = 100 > 25$ is sufficiently large) we have approximately:
$\sum_{i=1}^{100} X_i \overset{\text{CLT}}{\sim} \boldsymbol{N}\left(\mathbf{100} \cdot \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{100} \cdot \frac{\mathbf{1}}{\mathbf{4}}\right)$,
so $P(\sum_{i=1}^{100} X_i \leq 58) \overset{\text{CLT}}{\approx} \Phi\left(\frac{58-50}{\sqrt{25}}\right) = \Phi(1.6)$ (= 94.52%).
(Note that we did not apply continuity correction here (as many students tend to do): only if we transfer from discrete – binomial, Poisson- to normal. The X_i`s and $\sum_{i=1}^{100} X_i$ are continuous in this case)

9. **a**. We assume that the 100 process times are independent and all have the same (**unknown!**) distribution with (known) $\mu = 95$ and σ = 20 (or $\sigma^2 = 400$).
b. According to the CLT we have approximately: $\overline{X} = \frac{1}{100}\sum_{i=1}^{100} X_i \overset{\text{CLT}}{\sim} N\left(\mu, \frac{\sigma^2}{100}\right)$, so $N(95, 4)$ and
$P(\overline{X} > 100) \overset{\text{CLT}}{\approx} 1 - \Phi\left(\frac{100-95}{\sqrt{4}}\right) = 1 - \Phi(2.5) = 0.62\%$

10. **a**. $X \sim B(250, 0.25)$, so X is approximately $N(np, np(1-p)) = N(62.5, 46.5)$
(the rule of thumb for normal approximation: $n \geq 25$, $np > 5$ and $n(1-p) > 5$ is fulfilled)
b. "In reality the same as during last elections" implies $p = 0.25$: the distribution of X in a.
If we compute 22% of the sample size 250, we find 55 voters:
$P(X \leq 55) \overset{\text{c.c.}}{=} P(X \leq 55.5) \overset{\text{CLT}}{\approx} \Phi\left(\frac{55.5-62.5}{\sqrt{46.5}}\right) \approx \Phi(-1.03) = 0.1515$
c. Now $\mu = np = 250 \cdot 0.01 = 2.5 < 5$, so apply the **Poisson approximation** with μ = 2.5.
Applying the Poisson table we find: $P(X > 5) = 1 - P(X \leq 5) \approx 4.2\%$

11. **a**. $X \sim B(100, p)$, so μ = $100p$ and $\sigma^2 = 100p(1-p)$.
b. If X is approximately $N(100p, 100p(1-p))$, we find for $\frac{X}{100}$: $\frac{X}{100} \overset{\text{CLT}}{\sim} N\left(p, \frac{p(1-p)}{100}\right)$,
since $E\left(\frac{X}{100}\right) = \frac{1}{100}E(X) = \frac{1}{100} \cdot 100p = p$
and $var\left(\frac{X}{100}\right) = \frac{1}{100^2}var(X) = \frac{1}{100^2} \cdot 100p(1-p) = \frac{p(1-p)}{100}$, so:
$P\left(-0.05 \leq \frac{X}{100} - p \leq 0.05\right)$

$$= P\left(-\frac{0.05}{\sqrt{\frac{p(1-p)}{100}}} \leq Z \leq \frac{0.05}{\sqrt{\frac{p(1-p)}{100}}}\right) = \Phi\left(\frac{0.5}{\sqrt{p(1-p)}}\right) - \Phi\left(-\frac{0.5}{\sqrt{p(1-p)}}\right)$$

If we use that $p(1-p) \leq \frac{1}{4}$, This probability is at least $\Phi(1) - \Phi(-1) = 68.21\%$

12. X = "The demand of iPhones during 6 days" is Poisson($\mu = 6 \cdot 6 = 36$)-distributed.
This distribution can be normally approximated by a $N(\mu, \mu) = N(36, 36)$-distribution.
(*Note: as a rule of thumb one can use that if $\mu > 10$ the normal approximation of the Poisson distribution is allowed: for $\mu > 10$ no Poisson tables are available. Of course in practice one could compute the exact Poisson probabilities with Excel or a Graphical calculator).*
X is discrete (can only attain integer values): we can apply continuity correction for a more accurate normal approximation:

a. $P(X \leq 40) \overset{\text{c.c.}}{=} P(X \leq 40.5) = \Phi\left(\frac{40.5-36}{\sqrt{36}}\right) = \Phi(0.75) = 77.34\%$

b. If s is the desired (integer) number of iPhones in stock, then:

$P(X \leq s) \overset{\text{c.c.}}{=} P(X \leq s+0.5) = \Phi\left(\frac{s+0.5-36}{\sqrt{36}}\right) \geq 99\%$, if $\frac{s+0.5-36}{\sqrt{36}} = 2.33$

so $s = 35.5 + 2.33 \cdot 6 = 49.48$. The safety stock s should be (at least) 50.

13.

a. X is B(15, 0.01), so $P(X \geq 1) = 1 - P(X = 0) = 1 - 0.99^{15} = 14.0\%$

b. X is B(200, 0.01), so $\mu = 200 \times 0.01 = 2 < 10$, so X is appr. Poisson distr. with par. $\mu = 2$.

$P(X \leq 3) = 0.857$ (Poisson –table with $\mu = 2$)

c. X is B(4000, 0.01), so $E(X) = np = 4000 \cdot 0.01 = 40$ and $var(X) = np(1-p) = 39.6$

X is approximately normal with $\mu = 40 > 10$ and $\sigma = \sqrt{39.6} \approx 6.29$, so:

$P(X \geq 50) \overset{\text{c.c.}}{=} P(X \geq 49.5)$ (*continuity correction*)

$\approx P\left(Z \geq \frac{49.5-40}{6.29}\right) = P(Z \geq 1.51) = 1 - P(Z \leq 1.51) = 1 - 0.9345 \approx 6.5\%$