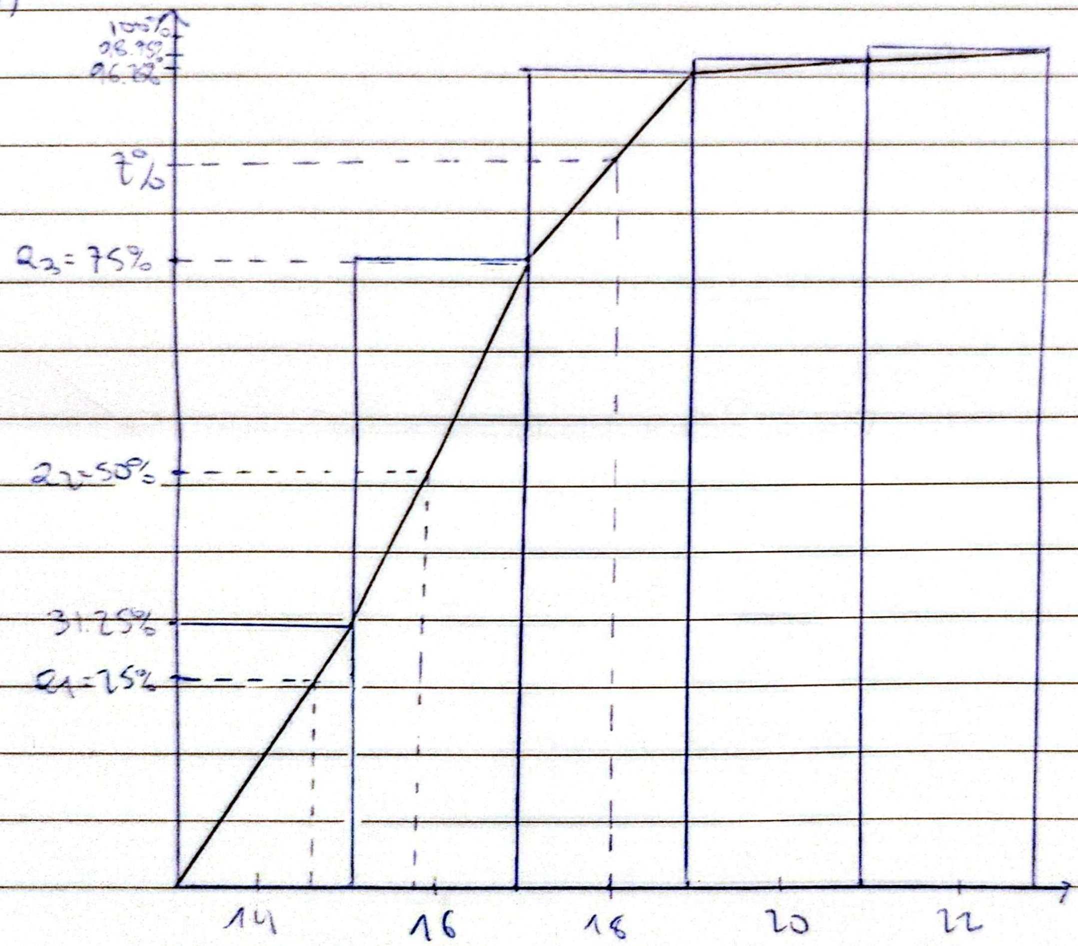
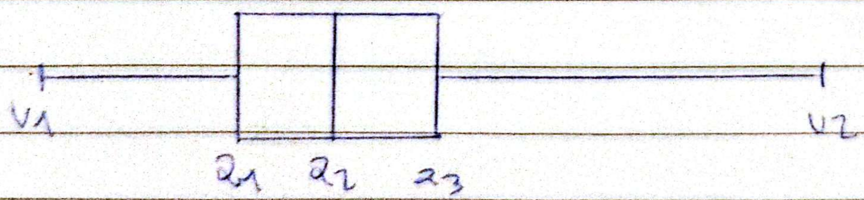
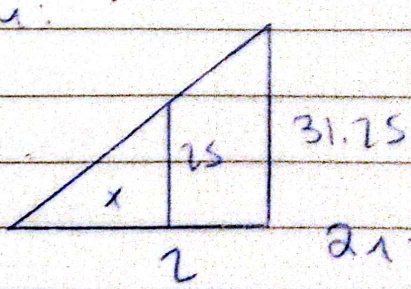
1-

b)





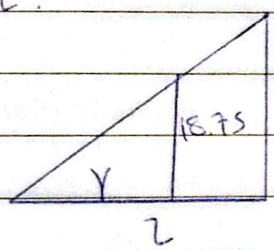
Q_1:



$$\frac{31.25}{2} = \frac{25}{x} \Rightarrow x = \frac{25 \cdot 2}{31.25} = 1.6$$

$Q_1 = 13 + x = 14.6$ days

Q2:



$$\frac{43.75}{2} = \frac{18.75}{y} \Rightarrow y = \frac{18.75 \cdot 2}{43.75} = 0.857$$

$$Q_2 = 15 + y = 15.857 \text{ days}$$

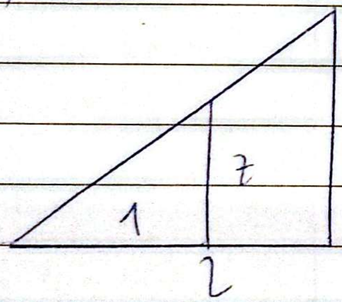
$Q_3 = 17$ days

~~$V_1 = Q_1 - 1.5\, IQR = 14.6 - 1.5(17 - 14.6) = 11$~~

$V_2 = Q_3 + 1.5\, IQR = 17 + 1.5(17 - 14.6) = 20.6$

There is an outlayer because $V_2 (20.6)$ is lower than the interval $[21, 23)$ in which there is one patient. So there is one outlayer in the interval $[21, 23)$

c)



$$\frac{21.25}{2} = \frac{z}{1} \Rightarrow z = \frac{21.25}{2} = 10.625$$

$$P_z = 75 + 10.625 = 85.625\%$$

$$100 - 85.625 = 14.375\%$$

14.375% of the sample should undergo this checkup

d)

$$\bar{x} = \frac{\sum n_i x_i}{N} = \frac{1278}{80} = 15.975 \text{ days}$$

$$s^2 = \frac{\sum n_i (x_i - \bar{x})^2}{N} = \frac{235.95}{80} = 2.95 \text{ days}^2 \rightarrow$$

$$g_1 = \frac{\sum n_i (x_i - \bar{x})^3}{N \cdot s^3} = \frac{297.6975}{80 \cdot (\sqrt{2.95})^3} = 0.7344$$

$$g_2 = \frac{\sum n_i (x_i - \bar{x})^4}{N \cdot s^4} - 3 = \frac{2508.8849}{80 \cdot (2.95)^2} = 0.6037$$

As $-2 < g_1 < 2$ and $-2 < g_2 < 2$, we could consider that this sample comes from a normal distribution

The shape of the sample is a little bit leptokurtic ($g_2 > 0$) and a little bit skewed to the left ($g_1 > 0$)

e)

$$CVx = \frac{Sx}{\bar{x}} = \frac{\sqrt{2.95}}{15.975} = 0.1075$$

$$\bar{y} = 1 + \frac{1}{2}\bar{x} = 1 + \frac{15.975}{2} \Rightarrow \bar{y} = 8.9875 \text{ days}$$

$$Sy = b \cdot Sx = \frac{1}{2}\sqrt{2.95} = 0.8588$$

$$CVy = \frac{Sy}{\bar{y}} = \frac{0.8588}{8.9875} = 0.0956$$

CVy is a little bit closer to 0 than CVx, so the recovery time with treatment would have a little bit more representative mean than the recovery time without treatment.

3:

$P(D) = 0.08 \quad P(I) = 0.15 \quad P(D|I) = 0.25$

a)

$$P(D|I) = \frac{P(D\cap I)}{P(I)} = 0.25 \Rightarrow P(D\cap I) = 0.25 \cdot P(I) =$$

$$= 0.25 \cdot 0.15 \Rightarrow P(D\cap I) = 0.0375$$

$$\left[P(D\cup I) = P(D) + P(I) - P(D\cap I) = 0.08 + 0.15 - 0.0375 = \right.$$
$$\left. = 0.1925\right]$$

b) $P(I\cap \bar{D}) = P(I) - P(I\cap D) = 0.15 - 0.0375 =$

$$\Rightarrow \left[P(I\cap \bar{D}) = 0.1125\right]$$

c)

$$P(D|\bar{I}) = \frac{P(D\cap \bar{I})}{P(\bar{I})} = \frac{P(D) - P(D\cap I)}{1 - P(I)} = \frac{0.08 - 0.0375}{1 - 0.15} \Rightarrow$$

$$\Rightarrow \left[P(D|\bar{I}) = 0.05\right]$$

d) $P(D\cap I) = 0.0375$

$P(D)\cdot P(I) = 0.08 \cdot 0.15 = 0.012$

$P(D\cap I) \neq P(D)\cdot P(I) \Rightarrow$ Dizziness and irritation can't be considered independent symptoms

2:

a) $\bar{x} = \frac{\sum x_i n_i}{N} = \frac{94}{4} = 23.5\,°C \quad \bar{y} = \frac{\sum n_i y_i}{N} = \frac{19}{4} = 4.75$ days

$$s^2_x = \frac{\sum n_i x_i^2}{N} - \bar{x}^2 = \frac{2316}{4} - (23.5)^2 = 26.75\,°C^2$$

$$s^2_y = \frac{\sum n_i y_i^2}{N} - \bar{y}^2 = \frac{109}{4} - (4.75)^2 = 4.6875 \text{ days}^2$$

$$s_{xy} = \frac{\sum x_i y_i n_i}{N} - \bar{x}\bar{y} = \frac{403}{4} - (23.5 \cdot 4.75) \Rightarrow$$

$$\Rightarrow s_{xy} = -10.875$$

b)

$$y = \bar{y} + \frac{s_{xy}}{s^2_x}(x - \bar{x}) \Rightarrow y = 4.75 + \frac{-10.875}{26.75}(x - 23.5) \Rightarrow$$

$$\Rightarrow y = 4.75 - 0.407x + 9.55 \Rightarrow y = -0.407x + 14.3$$

If $x = 12 \Rightarrow y = -0.407(12) + 14.3 \Rightarrow \boxed{y = 9.42 \text{ days}}$

c)

The slope in the previous equation is -0.407, this means that for each degree the temperature drops, the time will increase 0.407 days
~~decrease~~

d)

$$x = \bar{x} + \frac{s_{xy}}{s^2_y}(y - \bar{y}) \Rightarrow x = 23.5 + \frac{-10.875}{4.6875}(y - 4.75) \Rightarrow$$

$$\Rightarrow x = 23.5 - 2.32y + 11.02 \Rightarrow x = -2.32y + 34.52$$

$$\text{If } y = 3 \Rightarrow x = -2.32(3) + 34.52 \Rightarrow \boxed{x = 27.56\ °C}$$

e) The prediction in the question d) is more reliable than the prediction in the question b) because in our sample, we studied a temperature between 31 °C and 17 °C, and 12 °C is out of that range; whereas we studied a time of between 2 and 8 days, and the prediction in question d) (3 days) was in that range)

f)

a)

$$r^2 = \frac{s^2_{xy}}{s^2_x \cdot s^2_y} = \frac{(-10.875)^2}{26.75 \cdot 4.6875} = 0.9432$$

Strong relation between x and y

5.-

a) $p = 0.97 \rightarrow q = 0.03$, $n = 10$ $\Big] X \sim B(10, 0{,}97)$

$P(x \geq 8) = P(x=8) + P(x=9) + P(x=10)$

$P(x=8) = \binom{10}{8} \cdot 0.97^8 \cdot 0.03^2 = 0.0317$

$P(x=9) = \binom{10}{9} \cdot 0.97^9 \cdot 0.03 = 0.2281$

$P(x=10) = \binom{10}{10} \cdot 0.97^{10} = 0.7374$

$\left[P(x \geq 8) = 0.317 + 0.2281 + 0.7374 = 0.9972\right]_{//}$

b) $p = 0.03 \rightarrow X \sim B(200, 0.03) \Rightarrow X \sim P(\lambda)$

$\lambda = 200 \cdot 0.03 = 6$

$\Big[P(x \leq 4) = P(x=0) + P(x=1) + P(x=2) + P(x=3) + P(x=4) =$

$= e^{-6} \cdot \left(\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!} + \frac{6^4}{4!}\right) = 0.2851\Big]_{//}$

4.-

	C	$\bar{C}$
Test +	VP	FP
Test -	FN	VN

VP → 0,00686

FP → ~~0,00665~~ 0,04965

FN → 0,00014

VN → 0,94335

Prevalencia: $VP + FN = 0.007 \rightarrow FP + VN = 1 - 0.007 = 0.993$

Sensibilidad: $P(+|D) = \frac{VP}{VP+FN} = 0.98 \Rightarrow VP = 0.98(VP+FN) =$

$= 0.98 \cdot 0.007 \Rightarrow VP = 0{,}00686$

$VP + FN = 0.007 \Rightarrow FN = 0.007 - VP = 0.007 - 0.00686 \Rightarrow$
$\Rightarrow FN = 0.00014$

Specificity: $P(-|\bar{D}) = \frac{VN}{FP + VN} = 0.95 \Rightarrow VN = 0.95(FP + VN) =$
$= 0.95 \cdot 0.993 = 0.94335$

$FP + VN = 0.993 \Rightarrow FP = 0.993 - VN = 0.993 - 0.94335 \Rightarrow$
$\Rightarrow FP = \cancel{0.00665}$ 0,04965 Just typo, perfectly reasoning

a) This probability is called Positive predicted value (PPV)

$\left[PPV: P(D|+) = \frac{VP}{VP + FP} = \frac{0.00686}{0.00686 + \cancel{0.00665}} = \cancel{0.5078}\right]$ 0,121

b) This probability is called Negative Predicted Value (NPV)

$\left[NPV: P(\bar{D}|-) = \frac{VN}{FN + VN} = \frac{0,94335}{0,94335 + 0,00014} = 0.9999\right]$

c)

$[\text{Correct diagnosis: } VP + VN = 0.00686 + 0.94335 =$
$= 0.9502]$

perfect!

6-

$\left.\begin{array}{l} X \sim (N, 100) \\ Z \sim (0, 1) \end{array}\right]$ Standarization: $Z = \frac{X - N}{100}$

a) $\left.\begin{array}{l} P(X \leq 368) = P\left(Z < \frac{368 - N}{100}\right) = 0.75 \\ P(Z < 0.675) = 0.75 \end{array}\right] \frac{368 - N}{100} = 0.675 \Rightarrow$
$368 - N = 67,5 \Rightarrow$
$\Rightarrow [N = 300.5]$

b) $P(X > 444) = 1 - P(X < 444) = 1 - P\left(Z < \frac{444 - 300.5}{100}\right) =$
$= 1 - P(Z < 1.44) = 1 - 0.9251 \Rightarrow P(X > 444) = 0.0749$

c)
$P(172 < X < 444) = P(X < 444) - P(X < 172) =$
$= 0.9251 - P\left(Z < \frac{172 - 300.5}{100}\right) = 0.9251 - P(Z < -1.29) =$
$= 0.9251 - 0.0985 \Rightarrow [P(172 < X < 444) = 0.8266]$