Log 708 - Chapter 2 Solutions

Halvard Arntzen

2.1

We note that Z is continuous, so we can interchange $\leq$ and $<$ as we want. Answers from using the normal table:

a. 0.023

b. 0.999

c. 0.000 (Using R, more precisely: 0.0000425)

2.2

Regarding W we can use the formulas for mean and standard deviation in the compendium, and the fact that W becomes normal, allows us to use standardization to answer the questions.

We can note that in all three cases, the mean is the same (960), since this is not affected by correlations. The standard deviations on the other hand will differ, and it may be interesting to see how. Note that in all cases, question 2 is answered by the general formula

$$b = \mu_w + z_{0.05} \cdot \sigma_w = 960 + 1.645 \cdot \sigma_w \ ,$$

i.e. only σ_w will differ in the three cases. This can be seen if you solve the question of finding b such that

$P[W \leq b] = 0.95$ for a general normal variable with parameters μ_w, σ_w. Below we give the results of this in the three cases.

a. We get $\sigma_w = 97.6,\ b = 1121$.

b. Now $\sigma_w = 72.1,\ b = 1079$.

c. Now $\sigma_w = 57.3,\ b = 1054$.

We see that as correlation goes from 0.90 downwards through 0.40, 0 and eventually to -0.40, the standard deviation

σ_w decreases. Accordingly, the 5% worst case cost comes down from 1121 to 1054. By definition of W the risk comes from variation in both X and Y. In a case where ρ_{xy} is high positive, scenarios where both costs are high become relatively probable and total risk is high. When the correlation drops, we will more often see high X together with low Y and vice versa. More of the individual risk in X and Y is *cancelled out*, as the correlation drops.

2.3

a. $\mu_1 = 20, \sigma_1 = \sqrt{8} = 2.83$.

b. $\mu_2 = 20, \sigma_2 = 4$.

c. $\mu_3 = 0, \sigma_3 = \sqrt{8} = 2.83$. (Note that here $Y = X - Y = aX + bY$ with $a = 1, b = -1$)

d. $\mu_4 = 15, \sigma_4 = 2$. (You can treat 5 as a random variable with 0 variance).

e. $\mu_5 = 25, \sigma_5 = \sqrt{8}$. (This is the same as previous question, just replacing X by $X + Y$.)

2.4

a. The sample size is large enough ($n \geq 30$) to allow the use of so-called z-interval, i.e. we may use the formula
$$\bar{x} \pm z_{\alpha/2}\frac{S}{\sqrt{n}}.$$
The resulting interval is [16.51, 17.49].

b. We find $\hat{p} = 0.25$, we see that $n\hat{p}(1 - \hat{p}) = 75$ so we are safe when using the formula from the text. We use
$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
to get the interval [0.207, 0.292]

c. In case (i) ME would be smaller. (Because we divide by $\sqrt{n}$.) In case (ii) ME would be larger. (Because $z_{0.005} > z_{0.025}$.)

d. Apparently, a small number of flats takes exceptionally long time before sale. In such cases the median can be a better choice to describe typical characteristics. The mean can be misleading in the sense that almost all of the sample (and thus probably the population) lies below the mean. This obviously does not happen with the median.

2.5

a. Use the formula...
$$\bar{x} \pm z_{0.025}\frac{s}{\sqrt{n}} = 164 \pm 1.715.$$

b. As the value 160 is relatively far below the lower confidence limit for μ, it is highly unlikely that we have $\mu < 160$. It appears to be clear that the goal is not reached.

c. Using the standard formula (noting that the sample is by far large enough) we get
$$\hat{p} \pm z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.125 \pm 0.016$$

d. Here (i) is true, while (ii) is false.

2.6

a. Let X denote a weeks sale. Then we use that X is normal. Standardization leads to

$$P[X < 1050] = P[Z < 1.5] = 0.9332.$$

b. We need to find b such that $P[X > b] = 0.05$. Some calculations based on the standard normal leads to $b = \mu + z_{0.05}\sigma = 1064.5$

c. Let Y denote the profit for a given week. Then $Y = -200 + 1.10X$ is a linear function of X. It follows that

$$\mu_y = -200 + 1.10\mu_x = 790, \quad \sigma_y = 1.10\sigma_x = 110.$$

2.7

a. The loading time is a $N(\mu, \sigma)$ variable, so by the empirical rule,

$$\mu \pm 2\sigma$$

gives the approximate interval [3,9]. Using 1.96 rather than 2 gives a slightly more exact interval.

b. This is found by standardizing the normal variable T = loading time. The probability is 0.8413.