Solutions Tutorial 6

ec1030, Stats

Jacco Thijssen

1. Let X denote the exam score. Then the information given in the exercise tells us that $X \sim \mathrm{N}(63, 64)$. So, for the Z-transformation we have

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 63}{8} \sim \mathrm{N}(0, 1).$$

   (a) Using the table with cumulative probabilities for the $\mathrm{N}(0, 1)$ we find that

$$\begin{aligned} P(\{\text{student obtains a I}\}) &= P(X \geq 70) \\ &= P\left(Z \geq \frac{70 - 63}{8}\right) \\ &= P(Z \geq .88) \\ &= 1 - P(Z \leq .88) \\ &= 1 - F(.88) = 1 - .8106 = .1840. \end{aligned}$$

   (b) We want to find $P(X < 40)$. Using the table and the symmetry of the $\mathrm{N}(0, 1)$ distribution (draw a picture!) we get that

$$\begin{aligned} P(X < 40) &= P\left(Z < \frac{40 - 63}{8}\right) \\ &= P(Z < -2.88) = P(Z > 2.88) \\ &= 1 - P(Z \leq 2.88) = 1 - F(2.88) \\ &= 1 - .9980 = .0020. \end{aligned}$$

   (c) We are asked to find a value a such that $P(X \leq a) = .93$. The first step in solving this problem is to find a value z such that $P(Z \leq z) = .93$. From the table we find that $F(1.48) = .9306$. In other words, $z = 1.48$. Using the formula to transform Z-values into X-values we obtain that

$$x = \mu + z\sigma = 63 + 1.48 \cdot 8 = 74.8.$$

So, the probability that a student obtains a grade of at most 74.8 is .93.

(d) The first step in solving a problem is to find a value z such that $P(|Z| > z) = .10$, i.e. a value z for which

$$P(\{Z \text{ lies outside } [-z, z]\}) = P(Z < -z) + P(Z > z) = .10.$$

Because the $\mathrm{N}(0, 1)$ distribution is symmetric it holds that the area under the bell-curve to the left of $-z$ is equal to the area under the bell-curve to the right of z (draw a picture!). So,

$$P(Z < -z) = P(Z > z).$$

Therefore, if the sum of these probabilities has to be equal to .18, and both probabilities are equal we must have that each probability is equal to .18/2=.09. So we can choose one of them and use a table to find

$$\begin{aligned} & P(Z > z) = .05 \\ \Longleftrightarrow\ & 1 - P(Z \leq z) = .05 \\ \Longleftrightarrow\ & 1 - F(z) = .05 \\ \Longleftrightarrow\ & F(z) = .95 \\ \Longleftrightarrow\ & z = 1.65. \end{aligned}$$

So, 10% of Z-values lie outside the interval $[-1.65, 1.65]$. Turning these two values into X-values we obtain

$$x = \mu + z \cdot \sigma = 63 + (-1.65) \cdot 8 = 49.8,$$

and

$$x = \mu + z \cdot \sigma = 63 + 1.34 \cdot 8 = 76.2.$$

So, 18% of grades are below 49.8 or above 76.2. Note that this is, indeed, a symmetric interval around the mean of 63.

(e) This is the mirror image of the previous question. Here we are asked to find a symmetric interval such that the probability that X is *within* the interval equals a prescribed probability, in this case .64. So, again, the first step consists of looking at Z, which follows a N(0, 1) distribution. Now we are looking for a z such that $P(|Z| \leq z) = .64$, i.e. a z such that $P(-z < Z < z) = .64$. Because of the complement rule, this implies that

$$P(Z < -z) + P(Z > z) = 1 - .64 = .36,$$

and, again since N(0, 1) is symmetric,

$$P(Z < -z) = P(Z > z).$$

So, $P(Z > z)$ must be equal to $.36/2 = .18$. Applying the complement rule (again) and using the table we find that

$$\begin{aligned} & P(Z > z) = .18 \\ \Longleftrightarrow\ & 1 - P(Z \leq z) = .18 \\ \Longleftrightarrow\ & 1 - F(z) = .18 \\ \Longleftrightarrow\ & F(z) = .82 \\ \Longleftrightarrow\ & z = .92. \end{aligned}$$

So, the probability that Z is within the interval $[-.92, .92]$ is .64. Turning these two values into X-values we obtain

$$x = \mu + z \cdot \sigma = 63 + (-.92) \cdot 8 = 55.64,$$

and

$$x = \mu + z \cdot \sigma = 63 + .92 \cdot 8 = 70.36.$$

So, 64% of grades are between 55.64 and 70.36. Note that this is, indeed, a symmetric interval around the mean of 63.

(f) Judging from experience you could say that 18% I class grades could be realistic. However, one would usually see more than .2% fails, so this would indicate that the normal distribution is maybe not the correct model choice. In reality grades seem to be more skewed to the left.

2. (a) Assuming that the students represent a random sample (i.e. are independently drawn) we are essentially dealing with a binomial distribution. So, in this case, $X \sim \text{Bin}(5, .55)$.

(b) We are asked to compute $P(X \geq 4)$. Because X is a discrete random variable which can take values in $\{0, 1, 2, \ldots, 5\}$, this means that this probability is equal to the probability that exactly 4 or exactly 5 students vote for the candidate. Since the two events are mutually exclusive, the probability of the union is equal to the sum of the probabilities. So, using the formula for the binomial distribution,

$$p(x) = \binom{5}{x} (.55)^x (.45)^{5-x},$$

or a table we obtain that

$$\begin{aligned} P(X \geq 4) &= P(X = 4) + P(X = 5) \\ &= p(4) + p(5) \\ &= .2059 + .5003 = .2562. \end{aligned}$$

If you use a table you have to be careful as a success probability is not given for values larger than .5. You have to reformulate the problem as follows. A success with probability .55 means a failure with probability .45. So, the probability of at least 4 successes is the same as the probability of at most 1 failure. So,

$$P(X \geq 4) = p(0) + p(1),$$

where p is the distribution function of a $\text{Bin}(5, .45)$ distribution.

(c) The normal approximation tells us that, if $Z \sim \text{N}(0, 1)$, it holds that

$$P(a < X < b) \approx P\left(\frac{a - \text{E}(X)}{\sigma_X} < Z < \frac{b - \text{E}(X)}{\sigma_X} \right).$$

Since, for a binomial distribution with parameters $n = 5$ and $p = .55$ it holds that

$$\text{E}(X) = np = 5 \cdot .55 = 2.75,$$

and

$$\sigma_X = \sqrt{\text{var}(X)} = \sqrt{np(1-p)} = \sqrt{5 \cdot .55 \cdot .45} = .1314.$$

Using a table for the N(0, 1) we therefore find that

$$\begin{aligned} P(X \geq 4) &\approx P\left(Z \geq \frac{4 - 2.75}{.1314}\right) \\ &= P(Z \geq 1.12) = 1 - P(Z \leq 1.12) \\ &= 1 - F(1.12) = 1 - .8686 = .1314. \end{aligned}$$

This approximation is quite far off the true probability. This happens because n is not large enough for the normal distribution to closely resemble the binomial distribution. In particular,

$$np(1-p) = 1.238 < 9.$$

(d) Let X be the number of voters out of 50 voters who voted for your candidate. Then $X \sim \text{Bin}(50, .55)$. We are asked to compute the probability that the candidate gets a majority out of 50 votes, i.e. that $P(X > 25)$. We cannot use exact probabilities as n is too large, but the normal approximation can be applied since

$$np(1-p) = 50 \cdot .55 \cdot .45 = 12.375 > 9.$$

Since

$$\text{E}(X) = np = 50 \cdot .55 = 27.5,$$

and

$$\sigma_X = \sqrt{\text{var}(X)} = \sqrt{np(1-p)} = 3.518,$$

we find that (using symmetry):

$$\begin{aligned} P(X > 25) &\approx P\left(Z > \frac{25 - 27.5}{3.518}\right) \\ &= P(Z > -.71) = P(Z < .71) \\ &= F(.71) = .7611. \end{aligned}$$

(e) This is essentially the same question as the previous with slightly different numbers. Let X be the number of voters out of 500 voters who voted for your candidate. Then $X \sim \text{Bin}(500, .55)$. We are asked to compute the probability that the candidate gets a majority out of 500 votes, i.e. that $P(X > 250)$. The normal approximation can be applied since

$$np(1-p) = 500 \cdot .55 \cdot .45 = 123.75 > 9.$$

Since

$$\text{E}(X) = np = 500 \cdot .55 = 275,$$

and

$$\sigma_X = \sqrt{\text{var}(X)} = \sqrt{np(1-p)} = 11.124,$$

we find that (using symmetry):

$$\begin{aligned} P(X > 250) &\approx P\left(Z > \frac{250-275}{11.124}\right) \\ &= P(Z > -2.25) = P(Z < 2.25) \\ &= F(2.25) = .9878. \end{aligned}$$

(f) The larger the number of voters, the higher the probability that your candidate win. So I would want to get out the vote if I were you!

3. (a) You buy 5 units of X and sell 2 units of Y (i.e. you “buy” -2 units of Y). Therefore,

$$W = 5X - 2Y.$$

(b) The mean of the sum is the sum of the expectations, so

$$\begin{aligned} \text{E}(W) &= \text{E}(5X - 2Y) = 5\mu_X - 2\mu_Y \\ &= 5 \cdot 100 - 2 \cdot 200 = 100. \end{aligned}$$

Since X and Y are independent, the variance of the sum is the sum of the variances, so

$$\begin{aligned} \text{var}(W) &= \text{var}(5X - 2Y) = 5^2\sigma_X^2 + 2^2\sigma_Y^2 \\ &= 25 \cdot 100 + 4 \cdot 400 = 4100. \end{aligned}$$

(c) Regardless of whether X and Y are independent, the mean of the sum is the sum of the expectations, so

$$\begin{aligned} \mathrm{E}(W) &= \mathrm{E}(5X - 2Y) = 5\mu_X - 2\mu_Y \\ &= 5 \cdot 100 - 2 \cdot 200 = 100. \end{aligned}$$

Since X and Y are not independent, the variance of the sum is no longer the sum of the variances. We have to take the covariance into account, so

$$\begin{aligned} \mathrm{var}(W) &= \mathrm{var}(5X - 2Y) = 25\sigma_X^2 + 4\sigma_Y^2 - 2 \cdot 5 \cdot 2 \cdot \mathrm{cov}(X, Y) \\ &= 4100 - 2 \cdot 5 \cdot 2 \cdot (-40) = 4100 + 800 = 4900. \end{aligned}$$

(d) Due to the mistaken assumption of independence, the advisor's assessment of the variance is too low. Since the variance is an indication of the risk of the trading strategy, the advisor is underestimating the risk the client is exposed to.

4. Same as Question 1.

5. Same as Question 2.