# 475.101 / 102 / 107 / 108 Semester 2, 2000
# Assignment 3 Solutions

**Question 1.**

**(a)** **(i)** $\text{pr}(X \le 17) = 0.3694$

**(ii)** $\text{pr}(X < 17) = \text{pr}(X \le 17) = 0.3694$

**(iii)** $\text{pr}(X > 21) = 1 - \text{pr}(X \le 21) = 1 - 0.8413 = 0.1587$

**(iv)** $\text{pr}(12 \le X \le 16) = \text{pr}(X \le 16) - \text{pr}(X \le 12) = 0.2525 - 0.0228 = 0.2297$

**(b)** **(i)** The minimum amount of time that should be set aside so that the shop owner can organise the days banking without interrupting other work on 90% of days is 29 minutes.

**(ii)** The upper quartile for ice cream sales = 26.47 minutes.
The lower quartile for ice cream sales = 20.73 minutes.
The interquartile range for the ice cream sales = 26.47 – 20.73 = 5.74 minutes.

**(c)** **(i)** $\text{pr}(X > 6) = 1 - \text{pr}(X \le 6) = 1 - 0.0445 = 0.9555$

**(ii)** $\text{pr}(7.7 \le X \le 10.6) = \text{pr}(X \le 10.6) - \text{pr}(X \le 7.7)$
$= 0.9234 - 0.2931 = 0.6303$

**(iii)** $\text{pr}(X \le x) = 0.7$ when $x = 9.27$

$X \sim \text{Normal}(\mu = 8.5, \sigma = 1.47)$

| $x$ | $\text{pr}(X \le x)$ |
|---|---|
| 6 | 0.0445 |
| 7.7 | 0.2931 |
| 10.6 | 0.9234 |

| $\text{pr}(X \le x)$ | $x$ |
|---|---|
| 0.7 | 9.2709 |

**(d)** **(i)** $Y = 2X - 5W = 2 \times X + (-5) \times W \ \Rightarrow$

$E[Y] = E[2X-5W] = 2 \times E[X] + (-5) \times E[W] = 2 \times (-2) + (-5) \times 4 = -24$ and

$$\text{sd}[Y] = \text{sd}[2X-5W] = \sqrt{2^2 \times sd[X]^2 + (-5)^2 \times sd[W]^2}$$
$$= \sqrt{2^2 \times 5^2 + (-5)^2 \times 2^2}$$
$$= \sqrt{100+100} = \sqrt{200} \approx 14.14$$

**(ii)** We cannot say anything about the shape of the distribution of $Y$ as we don't know about the shape of the distributions of $X$ or $W$.

**Question 2.**

**(a)** The $z$-score for the statistics lecturers fish is $\dfrac{40-29.6}{9.5} = 1.09$.

The $z$-score for the mathematics lecturers fish is $\dfrac{42-38.4}{4.2} = 0.86$

Based on the $z$-scores, catching a perch that is a 40cm or longer is less likely to occur than catching a bream that is 42cm or longer. Therefore the statistics lecturers fish is a more impressive member of its species and so mathematics lecturer should pay for dinner.

**(b)** Let $P$ be the length of the perch. $P \sim \text{Normal}(29.6, 9.5)$
Let $B$ be the length of the bream. $B \sim \text{Normal}(38.4, 4.2)$
We want to know $\text{pr}(B > P) = \text{pr}(B - P > 0) = \text{pr}(D > 0)$

where $D = B - P \sim \text{Normal}(38.4 - 29.6 = 8.8,\ \sqrt{4.2^2 + 9.5^2} \approx 10.39)$.

From computer, $\text{pr}(D > 0) = 1 - \text{pr}(D < 0) = 1 - 0.1984 = 0.8016$.

There is approximately an 80% chance that the bream is longer than the perch.

It was assumed that the length of the two fish were independent of each other.

**Question 3.**

**(a)** $T = 17 \times C$ is assuming that all 17 accounts take exactly the same amount of time to check.

Let $C_i$ be the time taken to spot check the $i^{th}$ set of businesses accounts. $C_i \sim \text{Normal}(\mu = 15.8$ minutes, $\sigma = 3.4$ minutes) for each $i = 1,\ldots,17$

$C_{Tot} = C_1 + \ldots + C_{17}$

$E[C_{Tot}] = E[C_1] + \ldots + E[C_{17}] = 17E[C_i] = 268.6$

$\text{sd}[C_{Tot}] = \sqrt{sd[C_1]^2 + \ldots + sd[C_{17}]^2} = \sqrt{17}\,\text{sd}[C_i] \approx 14.02$

As the $C_i$ are all Normally distributed, so is $C_{Tot}$. So $C_{Tot} \sim \text{Normal}(268.6, 14.02)$.

**(b)** **(i)** From computer, $\text{pr}(C_{Tot} < 300) = 0.9874$. The inspector should complete 17 spot checks in the allocated time in nearly 99% of days.

**(ii)** $R = 300 - C_{Tot} = 300 + (-1) \times C_{Tot}$. As $C_{Tot}$ is Normally distributed, so is $R$.
$E[R] = 300 + (-1) \times E[C_{Tot}] = 300 - 268.6 = 31.4$.
$\text{sd}[R] = |-1| \times \text{sd}[C_{Tot}] = \text{sd}[C_{Tot}] = 14.02$.
$R_{Tot} \sim \text{Normal}(31.4, 14.02)$.

**(iii)** From computer, $\text{pr}(R > 12) = 1 - \text{pr}(R < 12) = 1 - 0.0832 = 0.9168$. There is roughly a 92% chance of the inspector starting an 18th spot check.

**(c)** **(i)** Each set of accounts is either audited or not. There is a fixed number of accounts spot checked. There is an estimated 10% chance for each set of accounts that they will be audited. Since accounts are sampled at random, the results of one spot check should be independent of the results of another spot check (assuming the inspectors mood doesn't influence who gets audited).
$A$ would be approximately Binomial with $n = 17$ and $p = 0.1$.

**(ii)** $E[A] = n \times p = 17 \times 0.1 = 1.7$. $\text{sd}[A] = \sqrt{np(1-p)} = \sqrt{17 \times 0.1 \times 0.9} = \sqrt{1.53} = 1.237$

**(iii)** Let $\overline{A}$ be the average number of accounts forwarded on for auditing each day from 50 working days. $E[\overline{A}] = E[A] = 1.7$. $\text{sd}[\overline{A}] = \dfrac{sd[A]}{\sqrt{n}} = \dfrac{1.237}{\sqrt{50}} = 0.1749$. As the sample size of 50 is large, by the Central Limit Theorem, the distribution of $\overline{A}$ will be approximately Normal.

**(iv)** As the distribution of $A$ is not Normal, the distribution of $\overline{A}$ will be approximately Normal.

**Question 4.**

**(a)** Let $X$ be the amount of rice in a packet. $X \sim \text{Normal}(\mu = 506 \text{ gm}, \sigma = 2.7 \text{ gm})$

```
Normal with mean = 506.000 and standard deviation = 2.70000

        x     P( X <= x)
 500.0000        0.0131
 512.0000        0.9869
```

$\text{pr}(500 \le X \le 512) = \text{pr}(X \le 512) - \text{pr}(X \le 500) = 0.9869 - 0.0131 = 0.9738$.
The managers requirement will be exceeded if the machines specifications are correct.

**(b) (i)** Let $X_i$ = be the weight of the $i^{\text{th}}$ packet of rice. $X_i \sim \text{Normal}(\mu = 506 \text{ gm}, \sigma = 2.7 \text{ gm})$
Let $\overline{X}$ be the mean weight from a sample of 25 packets of rice.
As $X_i$ is Normal (from specifications), $\overline{X}$ is also Normal.

$E[\overline{X}] = \mu_X = 506$, Remember, $\overline{X} = \dfrac{X_1 + X_2 + ... + X_{25}}{25} = \dfrac{1}{25}\sum_{i=1}^{25} X_i$

$\text{sd}[\overline{X}] = \dfrac{\sigma_X}{\sqrt{n}} = \dfrac{2.7}{\sqrt{25}} = 0.54.$

$\overline{X} \sim \text{Normal}(\mu = 506 \text{ gm}, \sigma = 0.54 \text{ gm})$

**(ii)** The Central Limit Theorem was not needed to be able to answer this since the distribution of the weight of a single packet of rice was Normal (from specifications), then the distribution of the mean weights of a sample of packets of rice must also be Normal.

**(iii)** The central 95% of a Normal distribution corresponds to 1.96 standard deviations either side of the mean. In this case: $506 - 1.96 \times 0.54$ to $506 + 1.96 \times 0.54 = 504.94$ to $507.06$. The central 95% of values of $\overline{X}$ should fall between 504.94 and 507.06 (when using a sample of size 25).

**(c) (i)** Assuming the specifications for the machine are correct, then $p$, the true proportion of packets of rice outside the 500 – 512 gram weight range, is $1 - 0.9738 = 0.0262$.

$\hat{P}$ is approximately Normal ( $\mu_{\hat{P}} = p = 0.0262$, $\sigma_{\hat{P}} = \sqrt{\dfrac{p(1-p)}{n}} = \sqrt{\dfrac{0.0262 \times 0.9738}{800}} = 0.0056$)

**(ii)** Yes, the Central Limit Theorem was needed.

**(iii)** The central 95% of a the distribution corresponds to $0.0262 - 1.96 \times 0.0056$ to $0.0262 + 1.96 \times 0.0056 = 0.0151$ to $0.0373$. The central 95% of values of $\hat{P}$ should fall between 0.0151 and 0.0373 (when using a sample of size 800).

**(d) (i)** <u>**Stem-and-leaf plot of rice weights**</u>

Units: 500 | 1 = 500.1 grams.

```
498 |  3
499 |  2
500 |  8
501 |  4  6  9
502 |  4  6  8  9
503 |  4  9  9
504 |  0  4  5
505 |  4  5
506 |  3
507 |  6
508 |
509 |
510 |  3  4
511 |  9
512 |
513 |
514 |
515 |  0
```

**(d) (ii)** The sample of weights does not have an exactly Normal distribution. The plot is not perfectly symmetric.

**(iii)** The data appear to be skewed to the right. Most the data is in the range 500-506, but there is a long upper tail with values going as high as 515.

**(iv)** It is not likely that the sample data has come from a Normal distribution.

**(e)** Assuming the specifications are correct, we would expect 95% of samples of size 25 to give a sample mean between 504.94 and 507.06 (from (b) above). Getting a sample mean of 504.5 would be unusual as it is outside of this range. If the specifications are correct, there is less than a 5% chance of this happening. Thus, from this, we have some doubts about the specifications of the machine.

**(f)** Assuming the specifications are correct, we would expect 95% of samples of size 800 to give a sample proportion of packets outside the 500 – 512 gram range between 0.0151 and 0.0373 (from (c) above). Getting a sample proportion of 26 out of 800 = 0.0325 would not be unusual as it is contained inside this range. Thus, from this, we have nothing against the specifications of the machine.

**(g)** From the stem-and-leaf plot we could see that there problems with the data coming from a Normal distribution as the sample data is moderately right skewed. The sample mean of 504.5 from the sample of 25 packets of rice was unusually low assuming the specifications were correct. The proportion of packets with weights outside the 500–512 gm range from the sample of 800 packets was not particularly unusual assuming the specifications were correct. Overall, there are reasons to doubt that the stated specifications for the machine were correct.