UNIVERSITY OF NEW SOUTH WALES
DEPARTMENT OF STATISTICS

MATH5856 Introduction to Statistics and Statistical Computations
Tutorial Problems week 9, Solutions

1. (a) Noting that $\sum_{i=1}^n X_i \sim N(25, 100)$,

$$\begin{aligned} P\left(\sum_{i=1}^n X_i > 30\right) &= P\left(\frac{\sum_{i=1}^n X_i - 25}{10} > \frac{30-25}{10}\right) \\ &= P(Z > 0.5) \quad \text{(where } Z \sim N(0,1)) \\ &= 0.3085 \end{aligned}$$

(b) Noting that $\overline{X} \sim N\left(1, \frac{4}{25}\right)$,

$$\begin{aligned} P(\overline{X} < 0) &= P\left(\frac{\overline{X}-1}{2/5} < \frac{0-1}{2/5}\right) \\ &= P(Z > 2.5) \quad \text{(where } Z \sim N(0,1)) \\ &= 0.0062 \end{aligned}$$

(c) Let $Z \sim N(0,1)$. Then

$$P(\overline{X} < k_1) = P\left(Z < \frac{k_1 - 1}{2/5}\right) = 0.3.$$

Hence

$$\frac{k_1 - 1}{2/5} = -0.525 \implies k_1 = -0.31.$$

Next, note that

$$\frac{\overline{X} - 1}{S/5} \sim t_{24}.$$

Then

$$P\left(\frac{\overline{X}-1}{S} < k_2\right) = P\left(\frac{\overline{X}-1}{S/5} < 5k_2\right) = 0.99$$

and so

$$5k_2 = t_{24,0.99} = 2.49 \implies k_2 = 0.498.$$

2. Let $X \sim N(30, 16)$. Then

$$\begin{aligned} P(20 < X < 40) &= P\left(\frac{20-30}{4} < \frac{X-30}{4} < \frac{40-30}{4}\right) \\ &= P(-2.5 < Z < 2.5) = 0.9876. \end{aligned}$$

So the probability that a single observation is between 20 and 40 is 0.9876. Let Y denote the number in a sample of 500 between 20 and 40. Then $Y \sim \text{Bin}(500, 0.9876)$ and

$$E(Y) = 500 \times 0.9876 = 493.8 \simeq 494.$$

So about 494 of the sample should be between 20 and 40.

3. All solutions can be found from tables for the chi-squared distribution or statistical software.

(a) If $Q \sim \chi^2_{11}$ then $P(Q < 3.05) = 0.01$.

(b) If $Q \sim \chi^2_{17}$ then $P(Q < 24.77) = 0.9$.

(c) If $Q \sim \chi^2_4$ then

$$P(Q > 0.484) = 1 - P(Q \leq 0.484) = 1 - 0.025 = 0.975.$$

(d) If $Q \sim \chi^2_{15}$ then

$$0.95 = P(Q < k) \implies k = \chi^2_{15,0.95} = 25.$$

4. Let $Q = \sum_{i=1}^n Z_i^2$. Then $Q \sim \chi^2_6$.

(a)

$$0.05 = P(Q > k) \implies k = \chi^2_{6,0.95} = 12.59.$$

(b)

$$P\left(\sum_{i=1}^n Z_i^2 > 16.81\right) = P(Q > 16.81) = 0.01.$$

5. First note that, since each $\frac{X_i-2}{5} \sim N(0,1)$ and are independent of each other,

$$Q_1 = \sum_{i=1}^{10} \left(\frac{X_i - 2}{5} \right)^2 \sim \chi^2_{10}.$$

Then

$$\begin{aligned} 0.01 &= P\left(\frac{1}{10} \sum_{i=1}^{n} (X_i - 2)^2 > a \right) \\ &= P\left(\frac{25}{10} \sum_{i=1}^{10} \left(\frac{X_i - 2}{5} \right)^2 > a \right) \\ &= P\left(\sum_{i=1}^{10} \left(\frac{X_i - 2}{5} \right)^2 > 10a/25 \right) \\ &= P(Q_1 > 10a/25). \end{aligned}$$

Hence,

$$10a/25 = \chi^2_{10,0.99} = 23.21 \implies a = 58.$$

Next, note that

$$Q_2 = \frac{9S^2}{\sigma^2} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{25} \sim \chi^2_9.$$

Then

$$\begin{aligned} 0.01 &= P\left(\frac{1}{9} \sum_{i=1}^{n} (X_i - \overline{X})^2 > b \right) \\ &= P\left(\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{25} > 9b/25 \right) \\ &= P(Q_2 > 9b/25). \end{aligned}$$

Hence,

$$9b/25 = \chi^2_{9,0.99} = 21.67 \implies b = 60.$$

6.

$$\begin{aligned}
P(\overline{X} > 1) &= P\left(\frac{\overline{X}-2}{2/5} > \frac{1-2}{2/5}\right), \quad \overline{X} \sim N\left(2, \frac{4}{25}\right) \\
&= P(Z > -2.5), \quad Z \sim N(0,1) \\
&= P(Z < 2.5) = 0.9938
\end{aligned}$$

Note that $\frac{\overline{X}-2}{S/5} \sim t_{24}$. Then

$$\begin{aligned}
0.95 &= P(\overline{X} < kS + 2) \\
&= P\left(\frac{\overline{X}-2}{S/5} < 5k\right) \\
&= P(T < 5k), \quad T \sim t_{24}
\end{aligned}$$

Hence,

$$5k = t_{24,0.95} = 1.71 \implies k = 0.34.$$

7. Each of the required distributions can be found using the results for the χ^2, t and F distributions.

(a)

$$\sum_{i=1}^{n} X_1^2 \sim \chi_n^2$$

(b)

$$\sum_{i=1}^{n} X_i \sim N(0,n) \implies \frac{\sum_{i=1}^{n} X_i}{\sqrt{\sum_{i=1}^{n} Y_i^2}} = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i}{\sqrt{\sum_{i=1}^{n} Y_i^2/n}} \sim t_n.$$

(c)

$$\frac{nX_1^2}{\sum_{i=1}^{n} Y_i^2} = \frac{X_1^2/1}{\sum_{i=1}^{n} Y_i^2/n} \sim F_{1,n}.$$

8. $0.9 = P(\sum_{i=1}^{4} Z_i^2 < a) = P(Q < a), \quad Q \sim \chi_4^2$

   $\Rightarrow a = \chi^2_{4,0.9} = 7.78.$

   $0.95 = P\left(\frac{(Z_1^2+Z_2^2)/2}{(Z_3^2+Z_4^4)/2} < b\right) = P(F < b), \quad F \sim F_{2,2}$

   $\Rightarrow b = F_{2,2,0.95} = 19.$

9. The fundamental results needed here are:

$$N_1 \sim N(\mu_1, \sigma_1^2),\ N_2 \sim N(\mu_2, \sigma_2^2) \quad \text{and independent}$$
$$\implies a_1 N_1 + a_2 N_2 \sim N(a_1\mu_1 + a_2\mu_2, b_1\sigma_1^2 + b_2\sigma_2^2)$$

   and

$$Z_1, \ldots, Z_n \overset{\text{ind.}}{\sim} N(0,1) \implies \sum_{i=1}^{n} Z_i^2 \sim \chi_n^2.$$

   Through repeated application of the first result we see that

$$\frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \tfrac{1}{n}).$$

   Next, note that $X_i - \mu \sim N(0,1)$ so application of the second result leads to

$$\sum_{i=1}^{n}(X_i - \mu)^2 + \sum_{i=1}^{n} Y_i^2 \sim \chi^2_{2n}.$$

   For the third, we could use the result established in the notes that says

$$(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$$

   for a general random sample from the $N(\mu, \sigma^2)$ distribution. In this instance $\sigma^2 = 1$. Noting that $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ it follows that

$$(n-1)S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 \sim \chi^2_{n-1}.$$