Statistics 5 Tutorial: Normal Distribution

Additional Practice Questions

1. 2011 Prelim/DHS/P2/Q9

In a certain junior college, the marks (out of 100) scored by a JC 1 student in a Class Test, Common Test and Promotional Examination are denoted by C, T and S respectively. C, T and S may be modelled by normal distributions with means and standard deviations as shown in the table below.

Type of assessment	Mean	Standard deviation
Class Test, C	68	α
Common Test, T	65	8
Promotional Examination, S	70	10

(i) Given that $\mathrm{P}(C > 85) = 0.05,$ determine the value of α.

In a particular year, a student sits for five Class Tests, a Common Test and a Promotional Examination. The average mark of the five Class Tests constitutes 20% of the overall assessment mark for the year. The Common Test and Promotional Examination constitute 20% and 60% of the overall assessment mark for the year respectively.

For the following parts, assume $\alpha = 10$.

(ii) Find the probability that the average mark scored by a student in the five Class Tests is more than 75.

(iii) Find the probability that a student scores an overall mark of more than 80 for the year.

(iv) State an assumption used in your calculations for (iii).

Solution:

(i)	$C \sim \mathrm{N}(68, \alpha^2)$ $T \sim \mathrm{N}(65, 8^2)$ $S \sim \mathrm{N}(70, 10^2)$ $\mathrm{P}(C > 85) = 0.05$ $\mathrm{P}\left(Z > \dfrac{85-68}{\alpha}\right) = 0.05$ $\dfrac{85-68}{\alpha} = 1.6449$ $\alpha = 10.335 = 10.3$ (3 s.f.)
(ii)	$C \sim \mathrm{N}(68, 10^2)$ $\overline{C} = \dfrac{C_1 + C_2 + \cdots + C_5}{5} \sim \mathrm{N}\left(68, \dfrac{10^2}{5}\right)$ $\mathrm{P}(\overline{C} > 75) = 0.058762 = 0.0588$ (3 s.f.)

(iii)	Let $X = 0.2\overline{C} + 0.2T + 0.6S$ $\mathrm{E}(X) = 0.2(68) + 0.2(65) + 0.6(70) = 68.6$ $\mathrm{Var}(X) = 0.2^2\left(\frac{10^2}{5}\right) + 0.2^2(8^2) + 0.6^2(10^2) = 39.36$ $X \sim \mathrm{N}(68.6, 39.36)$ $\mathrm{P}(X > 80) = 0.034601 = 0.0346 \quad (3 \text{ s.f.})$
(iv)	$\overline{C}$ or C, T and S are independent.

2. <u>2010 Prelim/CJC/P2/Q8</u>

The masses of snapper fish and pomfret fish sold by a fishmonger are normally distributed and independent of each other. The mean mass, standard deviation and selling price of snapper fish and pomfret fish are given in the following table:

	Snapper fish	Pomfret fish
Mean mass in kg	1	0.6
Standard deviation in kg	0.1	0.05
Selling price per kg in $	12	7

Find the probability that the

(i) total mass of 3 snapper fish and 2 pomfret fish is more than 4.5 kg;

(ii) mass of 3 snapper fish exceeds twice the mass of a pomfret fish by more than 1.85 kg;

(iii) total selling price of a snapper fish and 2 pomfret fish is more than $21.

A customer buys 15 fish, out of which n are snapper fish and the rest are pomfret fish. The probability that the customer pays more than $150 is less than 0.7. Find the largest value of n.

Solution:

(i)	Let X and Y be the r.v. mass of a snapper fish and a pomfret fish respectively. $X \sim N(1, 0.1^2)$; $Y \sim N(0.6, 0.05^2)$ $X_1 + X_2 + X_3 + Y_1 + Y_2 \sim N(4.2, 0.035)$ $P[X_1 + X_2 + X_3 + Y_1 + Y_2 > 4.5] = 0.0544$
(ii)	$X_1 + X_2 + X_3 - 2Y \sim N(1.8, 0.04)$ $P[X_1 + X_2 + X_3 - 2Y > 1.85] = 0.401$
(iii)	$12X + 7(Y_1 + Y_2) \sim N(20.4, 1.685)$ $P[12X + 7(Y_1 + Y_2) > 21] = 0.322$
	$12(X_1 + X_2 + \ldots + X_n) + 7(Y_1 + Y_2 + \ldots + Y_{15-n}) \sim N(63 + 7.8n, 1.8375 + 1.3175n)$ $P[12(X_1 + X_2 + \ldots + X_n) + 7(Y_1 + Y_2 + \ldots + Y_{15-n}) > 150] < 0.7$. Largest $n = 11$

3. 2011 Prelim/SAJC/P2/Q10

The operator of Queen Motorways records its weekly earnings from road toll charges according to the categories of vehicles using the road. The weekly earnings (in thousands of dollars) for each category are assumed to be normally distributed. These distributions are independent of one another and are summarised in the table below.

Vehicle Category	Mean (thousands)	Standard deviation (thousands)
Cars	120.3	10.4
Buses	69.2	12.5
Lorries	64.5	9.5

Find the probability that the difference in weekly earnings for buses and cars is not more than \$60,000.

Find the probability that over a 5-week period, the total earnings for lorries exceed \$345,000. What assumption must be made in your calculation?

Each week, the operator allocates part of the earnings for repairs. This is determined for each category of vehicle according to estimates of long-term damaged caused. It is calculated as follows:

x% of earnings from cars, 8% from buses and 15% from lorries.

Given that the probability that the total amount for repairs is at least \$25,000 in a given week is 0.097, find x.

Queen Motorways also records its weekly takings from collection of administration fees from drivers who pay the toll charges by cash. The mean weekly takings is \$2000 and the standard deviation of the weekly takings is \$800. State, with a reason, whether the weekly takings follows a normal distribution.

Solution:

	Let C and B be random variables "weekly earnings (in thousands of dollars) for cars and buses respectively". $C \sim N(120.3, 10.4^2)$ and $B \sim N(69.2, 12.5^2)$ $C - B \sim N(51.1, 10.4^2 + 12.5^2)$ $C - B \sim N(51.1, 264.41)$ $P(\lvert C - B \rvert \le 60) = P(-60 \le C - B \le 60) = 0.708$
	Let L be random variable "weekly earnings (in thousands of dollars) for lorries". $L \sim N(64.5, 9.5^2)$ $L_1 + L_2 + ... + L_5 \sim N(5(64.5), 5(9.5^2)) = N(322.5, 451.25)$ $P(L_1 + L_2 + ... + L_5 > 345) = 0.145$ The weekly earnings for lorries are independent of one another in the 5-week period.

	Let T be random variable "total amount for repairs (in thousands of dollars)". $T=\frac{x}{100}C+0.08B+0.15L$ $T\sim N(120.3\left(\frac{x}{100}\right)+15.211,108.16\left(\frac{x}{100}\right)^2+3.030625)$ $\mathrm{P}\left(T\geq 25\right)=0.097$ $P(T<25)=0.903$ $P(Z<\frac{25-E(T)}{\sqrt{Var(T)}})=0.903$ $\frac{9.789-120.3\frac{x}{100}}{\sqrt{108.16\left(\frac{x}{100}\right)^2+3.030625}}=1.2988$ From GC, $x=6.14$
	If a normal distribution is used to model the distribution of weekly takings, we would expect $\mathrm{P}\left(2-3(0.8)<X<2+3(0.8)\right)=0.998$ i.e. $\mathrm{P}\left(-0.4<X<4.4\right)=0.998$. However this is not reasonable as there is a significant range of weekly takings that are negative amounts. Hence the weekly takings should not follow a normal distribution. Alternatively, If $X\sim N\left(2000,800^2\right)$, $P\left(X<0\right)=0.00621$ (which is too large)

4. 2010 Prelim/TJC/P2/Q9

The times taken, in seconds, for two swimmers, *A* and *B*, to complete a 100-metre freestyle race are independent and normally distributed with means 48.0 and 47.2 and standard deviations 0.5 and 0.8 respectively. The two swimmers compete in a 100- metre race for which the world record is 46.9 seconds.

(i) Show that the probability that at least one of the two swimmers breaks the world record during the race is 0.363 correct to 3 significant figures.
(ii) Find the probability of Swimmer *B* beating Swimmer *A*.
(iii) If *A* and *B* are to meet 20 times for the 100m freestyle race, how many times do you expect *A* to beat *B*? Give your answer correct to the nearest integer.
(iv) Find the probability that the total sum of four randomly chosen timings of *A* is more than 4 times a randomly chosen timing of *B*.

Solution:

(i)	Let *A* be the timing of swimmer *A* for the 100m freestyle. $A \sim N(48.0, 0.5^2)$ Let *B* be the timing of swimmer *B* for the 100m freestyle. $B \sim N(47.2, 0.8^2)$ Required Probability $=1-$ P(non among the two broke the world record) $=1-P(A \geq 46.9)P(B \geq 46.9)$ $=1-(0.986097)(0.646170)=0.36281 \approx 0.363$
(ii)	$A-B \sim N(0.8, 0.89)$ $P(A>B)=P(A-B>0)=0.801781 \approx 0.802$
(iii)	$20 \times (1-0.801781) \approx 4$ Expected times out of 20 is 4 times.
(iv)	Let $W = A_1 + A_2 + A_3 + A_4 - 4B \sim N(3.2, 11.24)$ $P(A_1 + A_2 + A_3 + A_4 > 4B) = P(W>0) = 0.83008 \approx 0.830$

5. 2009/Prelim/RJC/P2/Q8

The random variable *X* is normally distributed with mean 20 and variance σ^2.

(i) Given that $\mathrm{P}(12 < X < k) = \mathrm{P}(X > k) = 0.42065$, show that σ = 8, correct to the nearest whole number, and find the value of the constant *k*.

X is related to a normal random variable *Y* by the formula $X = Y_1 + Y_2 - 4$, where Y_1 and Y_2 are two independent observations of *Y*. Another random variable *W*, where *W* and *Y* are independent, is normally distributed with mean 25 and variance 16.

(ii) Using $\sigma = 8$, find $\mathrm{E}(Y)$ and $\mathrm{Var}(Y)$, and hence find $\mathrm{P}(W > 2Y)$.

Solution:

(i)	Given $X \sim \mathrm{N}(20,\ \sigma^2)$. $\mathrm{P}(12 < X < k) = \mathrm{P}(X > k) = 0.42065$ $\Rightarrow \mathrm{P}(X \le 12) = 1 - 2(0.42065) = 0.1587$. $\Rightarrow \mathrm{P}(Z \le \frac{12-20}{\sigma}) = 0.1587$ where $Z \sim \mathrm{N}(0,\ 1)$ $\Rightarrow \frac{-8}{\sigma} = -0.99982$ (5 s.f.) (from graphic calculator) $\Rightarrow \sigma = \frac{-8}{-0.99982} = 8.0015$ (5 s.f.) Hence, $\sigma = 8$. (nearest whole number) (shown) □ $\mathrm{P}(X > k) = 0.42065$ $\Rightarrow \mathrm{P}(X \le k) = 1 - 0.42065 = 0.57935$ $\Rightarrow k = 21.6$. (3 s.f.) (from graphic calculator) □
(ii)	Using $\sigma = 8$, $X \sim \mathrm{N}(20,\ 8^2)$. $X = Y_1 + Y_2 - 4$ $\Rightarrow \mathrm{E}(X) = \mathrm{E}(Y_1 + Y_2 - 4)$ $\Rightarrow 20 = 2\,\mathrm{E}(Y) - 4$ $\Rightarrow \mathrm{E}(Y) = \frac{20+4}{2} = 12$. □ Also, $\mathrm{Var}(X) = \mathrm{Var}(Y_1 + Y_2 - 4)$ $\Rightarrow 64 = 2\,\mathrm{Var}(Y)$ (since Y_1 and Y_2 are independent) $\Rightarrow \mathrm{Var}(Y) = 32$. □ $\therefore\ Y \sim \mathrm{N}\left(12,\ \left(\sqrt{32}\right)^2\right)$. $\mathrm{E}(W - 2Y) = \mathrm{E}(W) - 2\mathrm{E}(Y) = 25 - 2(12) = 1$. $\mathrm{Var}(W - 2Y) = \mathrm{Var}(W) + 2^2\mathrm{Var}(Y) = 16 + 4(32) = 12^2$. $(W - 2Y) \sim \mathrm{N}(1,\ 12^2)$. $\mathrm{P}(W > 2Y) = \mathrm{P}(W - 2Y > 0) = 0.533$. (3 s.f.) □

6. The thickness, K cm, of a randomly chosen paperback book may be regarded as an observation from a normal distribution with mean 2 and variance 0.73.
The thickness, H cm, of a randomly chosen hardback book may be regarded as an observation from a normal distribution with mean 4.9 and variance 1.92.

(i) Find the probability that the combined thickness of 4 randomly chosen paperbacks is greater than the combined thickness of 2 randomly chosen hardbacks.

(ii) Find the probability that a randomly chosen paperback is less than half as thick as a randomly chosen hardback.

(iii) Find the probability that the difference in thickness between a randomly chosen hardback and the combined thickness of 2 randomly chosen paperbacks is more than 1 cm.

(iv) Find the probability that the average thickness of 10 hardback books is within 0.5 cm of its mean value.

Solution:

$K \sim N(2, 0.73) \qquad H \sim N(4.9, 1.92)$

i) To find: $P(K_1+K_2+K_3+K_4 > H_1+H_2)$

$\Rightarrow P(K_1+K_2+K_3+K_4-H_1-H_2 > 0)$

$E(K_1+K_2+K_3+K_4-H_1-H_2) = 4(2)-2(4.9) = -1.8$

$Var(K_1+K_2+K_3+K_4-H_1-H_2) = 4(0.73)+2(1.92) = 6.76$

$\therefore K_1+K_2+K_3+K_4-H_1-H_2 \sim N(-1.8, 6.76)$

$P(K_1+K_2+K_3+K_4 > H_1+H_2) = 0.2443719866$

$= 0.244$ #

ii) To find: $P(K < \frac{1}{2}H) \Rightarrow P(K-\frac{1}{2}H < 0)$

$K-\frac{1}{2}H \sim N(2-\frac{1}{2}(4.9), 0.73+\frac{1}{4}(1.92))$

$\sim N(-0.45, 1.21)$

$P(K<\frac{1}{2}H) = 0.6587634856$

$= 0.659$ #

iii) To find: $P(|H-(K_1+K_2)| > 1)$

$H-(K_1+K_2) \sim N(0.9, 3.38)$

$P(|H-(K_1+K_2)|>1) = P(H-(K_1+K_2)<-1)+P(H-(K_1+K_2)>1)$

$= 0.629$ #

iv) $\bar{H} = \frac{H_1 + \dots + H_{10}}{10} \sim N\left(4.9, \frac{1.92}{10}\right)$

$P(4.4 < \bar{H} < 5.4) = 0.746168688$

$= 0.746$ #

7. 2017/ACJC/Prelim II/Q10 (modified)

(a) An examination taken by a large number of students is marked out of a total score of 100. It is found that the mean is 73 marks and that the standard deviation is 15 marks. Give a reason why the normal distribution is not a good model for the distribution of marks for the examination. [1]

(b) The interquartile range of a distribution is the difference between the upper and lower quartile values for the distribution. The lower quartile value, l, of a distribution X, is such that $\mathrm{P}(X < l) = 0.25$. The upper quartile value, u, of the same distribution is such that $\mathrm{P}(X < u) = 0.75$.

The marks of another examination is known to follow a normal distribution. If a student who scores 51 marks is at the 80th percentile, and the interquartile range is found to be 10.8 marks, find the mean mark and the standard deviation of the marks scored by students who took the examination. [5]

(c) In a third examination, the marks scored by students are normally distributed with a mean of 52 marks and a standard deviation of 13 marks.

(i) If 50 is the passing mark and 289 students are expected to pass, how many candidates are there? [2]

(ii) Find the smallest integer value of m such that more than 90% of the candidates will score within m marks of the mean. [3]

(a)

Let X be the random variable 'marks of an examination'.

If $X \sim \mathrm{N}\left(73, 15^2\right)$, then **<u>99.7%</u>** of the values of X (the population) will lie within $73 \pm 3(15) = (28, 118)$ which contains a significant range of marks above 100. Hence, the normal distribution is not a suitable model.

<u>Alternatively,</u>

By GC, $\mathrm{P}(X > 100) = 0.0359$ if $X \sim \mathrm{N}(73, 15^2)$

i.e., there are significantly 3.59% of the students scoring more than the maximum mark of 100, which is impossible.

(b)

Let Y be the random variable 'marks of a school examination'.

$Y \sim \mathrm{N}(\mu, \sigma^2)$

$\text{P}(Y < 51) = 0.8$

$\text{P}(Z < \dfrac{51-\mu}{\sigma}) = 0.8$

$\dfrac{51-\mu}{\sigma} = 0.84162$

$\mu + 0.84162\sigma = 51$

$\text{P}(\mu - 5.4 < Y < \mu + 5.4) = 0.5$

$\text{P}(\dfrac{-5.4}{\sigma} < Z < \dfrac{5.4}{\sigma}) = 0.5$

$\text{P}(Z < -\dfrac{5.4}{\sigma}) = 0.25$

$-\dfrac{5.4}{\sigma} = -0.67449$

$\therefore \quad \sigma = 8.01$

$\therefore \quad \mu = 51 - 0.84162(8.0061) = 44.3$

(c) (i)

Let M be the random variable 'marks of another school examination'.

$M \sim \text{N}(52, 13^2)$

$\text{P}(50 < M) = 0.56113$

$\text{Number of passes} = (\text{total candidature}) \times 0.56113 = 289$

$\therefore$ total candidature $= 289 \div 0.56113 = 515$

(c) (ii)

$\text{P}(|M - 52| < m) > 0.9 \Rightarrow \text{P}(52 - m < M < 52 + m) > 0.9$ where $M \sim \text{N}(52, 13^2)$

$\Rightarrow \text{P}(M < 52 - m) < 0.05$

$\Rightarrow \qquad 52 - m < 30.6$

$\Rightarrow \qquad m > 21.4$

$\therefore$ Smallest integral value of $m = 22$

8. 2021/ASRJC/MYE/P2/Q4

A confectionary produces two types of cakes, strawberry and dark chocolate. The masses (in g) of strawberry cakes and dark chocolate cakes are modelled as having independent normal distributions with means and variances as shown in the table.

	Mean mass	Variance
Strawberry cakes	800	250
Dark chocolate cakes	600	σ^2

(i) Given that 90% of the dark chocolate cakes are more than 582 g, show that the value of σ is 14.045, correct to 3 decimal places. [2]

(ii) Three strawberry cakes are chosen at random. Find the probability that one has mass between 790g and 810g, while the other two have masses less than 790g. [2]

(iii) Find the probability that the total mass of three randomly chosen strawberry cakes differs from four times the mass of a randomly chosen dark chocolate cake by not more than 50 g. [3]

The cakes are sold by mass. The strawberry cake and dark chocolate cake are sold at \$4.52 and \$6.10 per 100g respectively.

(iv) Find, correct to the nearest dollar, the least value of *a* such that the probability that the total cost of one strawberry and two dark chocolate cakes purchased exceeding \$*a* is less than 0.2. [4]

(i)	Let *X* be the r.v. denoting "the mass of a strawberry cake." Let *Y* be the r.v. denoting "the mass of a dark chocolate cake." $X \sim \text{N}(800, 250) \qquad Y \sim \text{N}(600, \sigma^2)$ $P(Y > 582) = 0.90 \Rightarrow P(Z > \frac{582-600}{\sigma}) = 0.90$ $\Rightarrow -\frac{18}{\sigma} = -1.2815516 \Rightarrow \sigma = 14.04547 \Rightarrow \sigma = 14.045\,(3\text{ dp})$
(ii)	$3\,\text{P}(790 < X < 810)\,\text{P}(X < 790)^2 = 0.0985$ (3sf)
(iii)	Let $D = (X_1 + X_2 + X_3) - 4Y$, where $D \sim \text{N}\,(0,\ 3906.1924)$ $P(\lvert X_1 + X_2 + X_3 - 4Y \rvert \le 50)$ $= \text{P}(-50 \le X_1 + X_2 + X_3 - 4Y \le 50)$ $= 0.576$
(iv)	Let $T = \frac{4.52}{100}(X) + \frac{6.10}{100}(Y_1 + Y_2)$ $\text{E}(T) = 109.36$ $\text{Var}(T) = \left(\frac{4.52}{100}\right)^2 \text{Var}(X) + \left(\frac{6.10}{100}\right)^2 (2)\text{Var}(Y)$ $= 1.97878399$ $T \sim \text{N}(109.36, 1.97878399)$ $P(T > a) < 0.2$ $a > 110.54$ Least *a* = 111 (to the nearest dollar)