Statistics 5 Tutorial: Normal Distribution

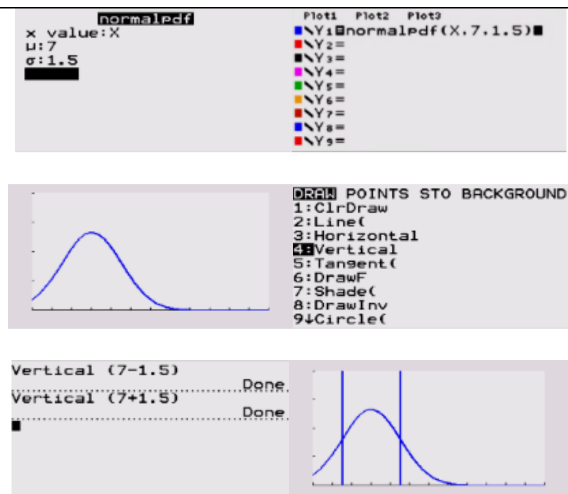
Basic Mastery Questions

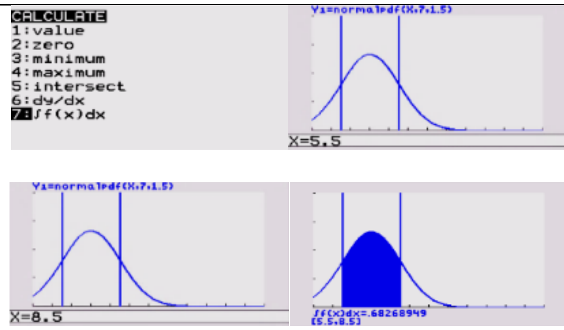
1. *Learning Experience Activity (20-30 mins)*
   Explore the characteristics of a normal curve: shape, centre, spread, the probability as area under the curve, and the empirical rule using a graphing tool.

   (1) **Shape, centre and spread**
   Carry out the following investigation:
   1. Start with a fixed μ, say 7.
   2. Study the behaviour of the normal curve as you change the values of σ, say from 1, 1.5, 2, or 2.5.
   3. Repeat (1) and (2) with different values of μ, say 8, 9 and 10.
   4. How is the shape of the normal curve affected by changing the mean? Changing the standard deviation?
   5. Explain why this behaviour is to be expected.

   (2) **Probability as area under the curve**
   Discuss: Can you think of an analogy why probability that a value lies between two values $x = a$ and $x = b$ is given by the area under the curve from $x = a$ and $x = b$?

   (3) **The empirical rule** (ref pt. 6 of section 5.2.1)
   Sketch a normal curve using any of the values of μ and σ given in Q(1), and find the percentage of the total area which is within:
   (a) 1 standard deviation of the mean
   (b) 2 standard deviations of the mean
   (c) 3 standard deviations of the mean





(1) The mean μ controls the location of the peak of the curve. Varying it will move the entire distribution to the left or right. On the other hand, the standard deviation σ controls whether the curve is broad and flat (larger standard deviation) or narrow and tall (smaller standard deviation). Varying it will not move the distribution left or right, but will affect the spread of the distribution.

(2) Use the analogy of a dartboard where the probability that a dart falls inside a defined compound is given by the ratio of the area of the defined compound to the total area of the dartboard.

(3) (a) 68.3% , (b) 95.4%, (c) 99.7%
(The <u>empirical rule</u> or the <u>68-95-99.7 rule</u>: Practically all of the population (99.7%) lies in the interval $\mu \pm 3\sigma$, about 95% of the population lies in the interval $\mu \pm 2\sigma$, and about 68% of the population lies in the interval $\mu \pm \sigma$.)

2. Given X and Y are independent with $X \sim \mathrm{N}(3,4)$ and $Y \sim \mathrm{N}(6,16)$. State the distribution of:

(a) $2X$,

(b) $X_1 + X_2$, where X_1 and X_2 are two independent observations of X,

(c) $X - 2Y$,

(d) $\dfrac{Y_1+Y_2+Y_3}{3}$, where Y_1, Y_2 and Y_3 are three independent observations of Y.

(a) $2X \sim \mathrm{N}(6,16)$

(b) $X_1 + X_2 \sim \mathrm{N}(6,8)$

(c) $X - 2Y \sim \mathrm{N}(-9,68)$

(d) $\dfrac{Y_1+Y_2+Y_3}{3} \sim \mathrm{N}\left(6, \dfrac{16}{3}\right)$

Why must X and Y be independent?

If X and Y are not independent,
$\mathrm{Var}(X-2Y) \neq \mathrm{Var}(X) + \mathrm{Var}(2Y)$

3. Given $X \sim \mathrm{N}(144, 49)$, find a if $\mathrm{P}(X > a) = 0.1182$.

$\mathrm{P}(X > a) = 0.1182$ Using GC (invNorm) $a = 152$

4. The independent random variables R and S each have normal distributions. The means of R and S are 10 and 12 respectively, and the variances are 9 and 16 respectively.
Find the following probabilities

(i) $\mathrm{P}(R < 12)$,

(ii) $\mathrm{P}(\overline{R} < 12)$, where $\overline{R}$ is the mean of a sample of 4 independent observations of R,

(iii) $\mathrm{P}(R < S)$,

(iv) $\mathrm{P}(2R > S_1 + S_2)$, where S_1 and S_2 are two independent observations of S.

(i)	$\mathrm{P}(R < 12) = 0.748$
(ii)	$\overline{R} \sim \mathrm{N}\left(10, \frac{9}{4}\right)$ $\mathrm{P}(\overline{R} < 12) = 0.909$
(iii)	$R - S \sim \mathrm{N}(-2, 25)$ $\mathrm{P}(R < S) = \mathrm{P}(R - S < 0) = 0.655$
(iv)	$2R - S_1 - S_2 \sim \mathrm{N}(-4, 68)$ $\mathrm{P}(2R > S_1 + S_2) = \mathrm{P}(2R - S_1 - S_2 > 0) = 0.314$

5. Given that $X \sim \mathrm{N}(70, \sigma^2)$, find the value of σ such that $\mathrm{P}(|X| > 102) = 0.2$.

Solution:

$$\mathrm{P}(|X| > 102) = 0.2$$
$$\mathrm{P}\left(\left|\frac{X-70}{\sigma}\right| > \frac{32}{\sigma}\right) = 0.2$$
$$\mathrm{P}\left(|Z| > \frac{32}{\sigma}\right) = 0.2$$
$$\mathrm{P}\left(Z < -\frac{32}{\sigma}\right) + \mathrm{P}\left(Z > \frac{32}{\sigma}\right) = 0.2$$
$$\mathrm{P}\left(Z < -\frac{32}{\sigma}\right) = 0.1 \quad \left(\text{or } \mathrm{P}\left(Z > \frac{32}{\sigma}\right) = 0.1\right)$$
$$-\frac{32}{\sigma} = -1.28155$$
$$\sigma = 25.0$$