<u>Normal Distribution Summary</u>

If X is a random variable which is **normally** distributed with mean μ and variance σ^2, then we say that $X \sim N(\mu,\ \sigma^2)$.

Note: <u>Any well-formed probability distribution function will have a mean and a variance</u> (eg Binomial and Poisson distributions both have their own means and variances), **so do not assert** that a distribution is normal in nature simply because the mean and variance values are given, UNLESS things have been stated explicitly within the question or sensible assumptions are required to be made.

Possible question types:

1. Mean and variances of the normal distribution are given, and a probability is to be calculated for a specific scenario (by far the easiest):

Eg: The mass of sugar in a 1kg bag may be assumed to have a normal distribution with mean 1005g and standard deviation 2g. Find the probability that a 1kg bag will contain less than 1000g of sugar.

<u>SOLUTION:</u>

Let the random variables X denote the mass of sugar in 1kg labelled bags.

Then $X \sim N(1005,\ 2^2)$ and $P(X < 1000) = 0.00621$ (shown)

2. Addition/Subtraction of multiple independent normal variables, in which case the following formulations are relevant:

Let X and Y be two independent normal distributions with mean/variance $\mu_1 / \sigma_1{}^2$ and $\mu_2 / \sigma_2{}^2$ respectively. Then $X \pm Y \sim N(\mu_1 \pm \mu_2,\ \sigma_1{}^2 + \sigma_2{}^2)$. Understand that this can be applied to three or more variables.

Eg: Weights of persons using a certain lift are normally distributed with mean 70kg and standard deviation 10kg. The lift has a maximum permissible load of 300kg. If 4 persons from this population are in the lift, determine the probability that the maximum load is exceeded.

SOLUTION :

Let the random variable X denote the weight of a person.

Then $X \sim N(70, 10^2)$, $T = X_1 + X_2 + X_3 + X_4 \sim N(280, 400)$

$P(T > 300) = 0.159$ (shown)

3. Multiples of a single variable are considered within the context of the question.

Eg. Using parameters from the lift problem previously, if one person from this population is in the lift and he has luggage weighing **3 times** his own weight, determine the probability that the maximum load is exceeded. (be extremely mindful of the above phrasing)

SOLUTION:

$4X \sim N(4\times 70 = 280, 4^2 \times 100 = 1600)$

$P(X + 3X > 300) = P(4X > 300) = 0.309$ (shown)

*4. Pricing of items are brought in as separate attributes.

Eg. Melons are sold by weight at a price of $1.50 per kilogram. The masses of melons are normally distributed with a mean of 0.8kg and a standard deviation of 0.1kg. Pumpkins are sold by weight at a price of $0.50 per kilogram. The masses of pumpkins are normally distributed with a mean of 1.2kg and a standard deviation of 0.2kg. Find the probability that the total price of 5 randomly chosen melons and 3 randomly chosen pumpkins exceeds $8.

SOLUTION :

Let X be the random variable denoting the price of a randomly chosen melon.

Then $X \sim N(0.8 \times 1.5 = 1.2,\ (0.1 \times 1.5)^2 = 0.15^2)$

Let Y be the random variable denoting the price of a randomly chosen pumpkin.

Then $Y \sim N(0.5 \times 1.2 = 0.6,\ (0.5 \times 0.2)^2 = 0.1^2)$

Let $T = (X_1 + X_2 + X_3 + X_4 + X_5) + (Y_1 + Y_2 + Y_3)$,

Then $T \sim N(78,\ 0.1425)$

$P(T > 8) = 0.298$ (shown)

> **5. Standardization of the normal variable(s) is required. Typically the inverse normal function will have to be employed.

Eg. A jam manufacturer produces a pack consisting of 8 assorted pots of differing flavours. The actual weight of each pot may be taken to have an independent normal distribution with mean 52g and standard deviation σ g. Find the value of σ such that 99% of packs weigh above 400g.

SOLUTION:

Let X be the random variable denoting the weight of a randomly chosen pot.

Then $X \sim N(52,\ \sigma^2)$

Let $T = X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8$

Then $T \sim N(416,\ 8\sigma^2)$

$P(T > 400) = 0.99 \Rightarrow P(T < 400) = 0.01$

$$P\left(Z < \frac{400-416}{\sqrt{8\sigma^2}}\right) = 0.01$$

$$\therefore \frac{400-416}{\sqrt{8\sigma^2}} = invNorm(0.01) = -2.326 \Rightarrow \sigma = 2.432 \text{ (shown)}$$

Approximations to the normal distribution:

Binomial to Normal: $np > 5, \quad nq > 5, \ n > 30 \ \rightarrow \ X \sim N(np, npq)$

Poisson to Normal: $\lambda > 10 \rightarrow \ X \sim N(\lambda, \lambda)$

Note that since both approximations involve discrete to continuous distributions, **continuity corrections** must be included to maintain accuracy of answers. This works as follows:

(i) $P(X > a) = P(X > a + 0.5)$

(ii) $P(X \geq a) = P(X > a - 0.5)$

(iii) $P(X < a) = P(X < a - 0.5)$

(iv) $P(X \leq a) = P(X > a + 0.5)$