Stats 1 Formula

Measure of location: $\bar{x}$ vs μ

Sample	Population
$\bar{x} = \frac{\sum x_i}{n}$	$\mu = \frac{\sum x_i}{n}$
$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$	$\mu = \frac{\sum x_i f_i}{\sum f_i}$
$\bar{x} = \frac{\sum (x_i - a)}{n} + a$	$\mu = \frac{\sum (x_i - a)}{n} + a$

Measure of spread:

Sample	Population
$s^2 = \frac{1}{n-1} \sum (x - \bar{x})^2$	$\sigma^2 = \frac{1}{n} \sum (x - \mu)^2$
$s^2 = \frac{1}{n-1} \left\{ \sum x^2 - \frac{(\sum x)^2}{n} \right\}$	$\sigma^2 = \frac{1}{n} \left(\sum x^2 \right) - \mu^2$
$s^2 = \frac{1}{n-1} \left\{ \sum fx^2 - \frac{(\sum fx)^2}{n} \right\}$	$\sigma^2 = \frac{1}{n} \left(\sum fx^2 \right) - \left\{ \frac{(\sum fx)}{n} \right\}^2$

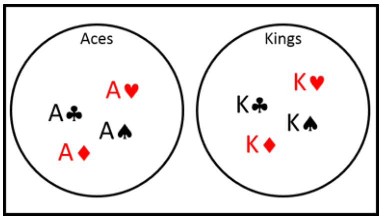
$$SD = \sigma = \sqrt{\text{var}} \quad \Leftrightarrow \quad \text{var} = \sigma^2$$

1 standard deviation includes 2/3 of data
2 standard deviations include 95% of data
3 standard deviations include 'almost all' data

"a quantity expressing by how much the members of a group differ from the mean value for the group"

Probability

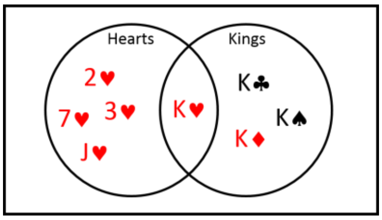
Mutually Exclusive



$$P(A \cup B) = P(A) + P(B)$$
$$P(A \cap B) = P(A) \times P(B)$$

$$P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$$

Non-Mutually Exclusive



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cap B) = P(A \mid B) \times P(B)$$
$$\Leftrightarrow$$
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Exclusive – no overlap of possible events
Exhaustive – all possible events are accounted for

If *any* of the following are true then they are *all* true & *events are independent*:

$$P(A \mid B) = P(A \mid B')$$
$$P(A \mid B) = P(A)$$
$$P(A \mid B') = P(A)$$

$$P(B \mid A) = P(B \mid A')$$
$$P(B \mid A) = P(B)$$
$$P(B \mid A') = P(B)$$

$$P(A \cap B) = P(A) \times P(B)$$

Binomial distribution (a discrete distribution):

$$X \sim B(n, p) \qquad P(X = x) = {}^{n}C_{x}\, p^{x} q^{n-x}$$

Expectation & Variance of Binomial distribution:

$$X \sim B(n, p)$$

$$E(x) = \mu = np$$
$$Var(x) = \sigma^2 = npq$$

Normal distribution (a continuous distribution):

$$Z \sim N(\mu, \sigma^2)$$

The standard normal distribution: $Z \sim N(0\,,1)$

Conversion: $Z = \frac{X-\mu}{\sigma}$

To find μ and/or σ using given probabilities:

1. Draw a sketch or curve
2. Use the probabilities table to find corresponding z value
3. Use the z value together with the conversion to find μ and/or σ. Finding both μ and σ will involve solving simultaneous equations.

Estimation

Sample mean = $\bar{x}$	Sample variance = s^2
Unbiased estimator of population mean = $\bar{X}$ (ie the distribution of many sample means)	Unbiased estimator of population variance = $\bar{S}$ (ie the distribution of many sample variances)

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\bar{Z} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

Confidence Interval	90%	95%	98%	99%	99.8%
z	1.6449	1.9600	2.3263	2.5758	3.0902

$$\bar{x} \pm \text{"z"} \frac{\sigma}{\sqrt{n}}$$

Product Moment Correlation Coefficient:

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

$$S_{xy} = \sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i$$

$$S_{xx} = \sum x_i^{\ 2} - \frac{1}{n}\left(\sum x_i\right)^2 \qquad S_{yy} = \sum y_i^{\ 2} - \frac{1}{n}\left(\sum y_i\right)^2$$

Regression:

y on x	x on y
$y = a + bx$	$x = a' + b'y$
$a = \bar{y} - b\bar{x}$	$a' = \bar{x} - b'\bar{y}$
$b = \frac{S_{xy}}{S_{xx}}$	$b' = \frac{S_{xy}}{S_{yy}}$