# Probability distributions

Use corresponding distribution models if the parameters in the model are present

### Calculation workings

*All calculations are assumed to be “less than” in workings*

Principle for discrete distributions:

Principle for normal distribution:

## Normal distribution [Continuous]

### Important properties

* Symmetrical
* Total area = 1

### Standardizing normal distribution to Z

|  |  |
| --- | --- |
| Binomial distribution [Discrete]Experiment conditions  1. Finite number of trials n 2. Independent outcomes 3. True of false results 4. Probability of success is constant   **When answering remember to be specific about what the outcomes are** Values | Poisson distribution [Discrete]Experiment conditions Events occur:   1. Singly 2. Randomly 3. Independently  Adjusting mean number Scale with the ratio of the different time intervals |

## Approximations

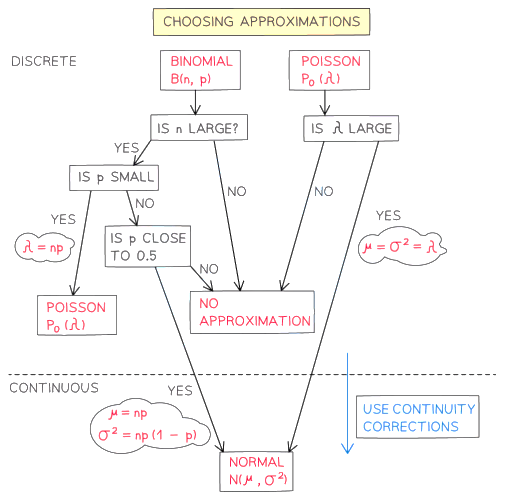
The question will let you know when to use approximations.

### Continuity corrections

It corrects differences between a discrete and continuous distribution when approximating.

So if approximating B with Po then correction is not needed as both are discrete.

**To apply continuity corrections:**

1. Imagine a discrete and uniform distribution on each other
2. Each bar has a width of 1 and each integer point is at the middle of each bar
3. Think if each “chonky bar” has to be included in the inequality ()

**Examples**

### Choosing approximation workflow

# Continuous random variables

**Example:**

Probability distribution function (p.d.f) 🡺 f(x)

**Example:**

Cumulative distribution function (c.d.f) 🡺 F(x)

Mode = Highest point of graph, solving by plotting and / or

Quartiles . For multiple functions in , find where quartiles are located first.

### Continuous uniform distribution

# Hypothesis testing

## Sampling & statistic distributions

Census - Measure every member of population Completely accurate

Sample - Measure a subset of population Less accurate

Population - Whole set of items that are of interest

Statistic - Quantity calculated only from observations in sample

is a statistic but isn’t

Sampling frame - A **list** of items

Sampling distribution - **Probability distribution** of the events

**When answering remember to be specific about what the items / events are**

## Test values

Test statistic - Result that is calculated from sample

Null hypothesis H0 - Hypothesis assumed correct

Alternative hypothesis H1 - Parameter if assumption shown wrong

## Types of tests

H1 in the form

H1 in the form

Over / under / increase / decrease = 1 tailed

Change / not = 2 tailed

## Critical values

Critical region - the range of values for the statistic that would reject H0

### Example – Finding critical values

**Conclusion**

For two tailed tests there are twocritical regions at each ends of the distribution. a.s.l. = sum of c.r.

## Carrying tests

### Identify the available data

1. Test statistic, X
2. Significance level
3. One tail or two tail

### Execution

1. Define X – the test statistic
2. Define H0 & H1
3. Assume H0 true, substitute p to distribution modelfind
4. Compare: if
5. Conclude in context of question

### Example

A seed producer claims that 96% of its beans turn golden

A random sample of 75 bean seeds was planted and 66 of these seeds turned golden

Test at 1% significance, whether of the producer is overstating the probability of the seeds turning golden

1. **Data identified**

* Model hence is Binomial
* Test statistic value
* s.l. = 0.01
* 1 tail

1. **Define X – the test statistic**

Let X = the number of seeds that turn golden

1. **Define H0 & H1**

H0: , H1:

1. **Assume H0 true, substitute p to distribution modelfind**
2. **Compare: if**

H0 rejected

1. **Conclude in context of question**

Producer is overstating …

### Example if two tailed:

Test at 1% significance, whether of the producer is lying about the probability of the seeds turning golden

1. **Define X – the test statistic**

Let X = the number of seeds that turn golden

1. **Define H0 & H1**

H0: , H1:

1. **Assume H0 true, substitute p to distribution modelfind & if two tailed**
2. **Compare: if**

H0 rejected (first tail)

1. **Conclude in context of question**

Producer is lying and overstating…