Modelling

Distribution	Conditions	Assumptions
Uniform		
Bernoulli		
Binomial		
Geometric		
Hypergeometric		

Example

The random variable C takes integral values in the interval -5 to 5, with probabilities

$$\mathrm{P}(C=-5)=\mathrm{P}(C=5)=\frac{1}{20},$$
$$\mathrm{P}(C=i)=\frac{1}{10},\quad \text{for } -4\leqslant i\leqslant 4.$$

Calculate the expectation and variance of C.

A shopper buys 36 items at random in a supermarket and, instead of adding up her bill exactly, she first rounds the cost of each item to the nearest 10 pence, rounding up or down with equal probability if there is an odd amount of 5 pence. Should she suspect a mistake if the cashier asks her for 20 pence more than she estimated? Explain your reasoning briefly but clearly.

Hypothesis Testing

Hypothesis testing is a process of decision-making when presented with data. The process consists of several steps. For example, suppose a darts player hits a triple 20 80% of attempts. He is learning a new stance and tries 20 throws. Let X be the number of triple 20s he throws using the new stance, then $X \sim B(20, p)$.

1. Setting up a **null hypothesis**, H_0. Typically based on past experience. In our example, the null hypothesis might be $p = 0.8$
2. Setting up an **alternative hypothesis**, H_1. There are multiple possible alternative hypotheses. For example, $p > 0.8$ (the player has improved), $p < 0.8$ (the player has got worse), $p \neq 0.8$ (the player's performance has changed).
3. Decide on a **significance level**. This is a measure of how unlikely it is that we reject our null hypothesis if it is true. Typical values used are 5% and 1%.
4. Find out the range of values where we would reject our null hypothesis. This range of values is the **critical region**. Depending on our alternative hypothesis, this can be **one-tailed** or **two-tailed**. The region where we accept the null hypothesis is called the **acceptance region**
5. Calculate our **test statistic**. We may already know it (for example, we might be told the number of successes), or we might have to calculate it, (for example we might need to calculate the average of the values we have measured).
6. Decide the outcome of the test based on whether or not the test statistic is inside the critical region or not. Sometimes it is necessary to give the exact value, the **p-value**, of the probability that we obtain a test statistic as extreme as we observed assuming H_0.

Binomial Hypothesis Testing

Example

A coin is tossed 20 times. The result is 14 heads and 6 tails. At the 10% significance level, determine the critical region and test whether the coin is biased.

Let X be ____. Then $X \sim$ ______ where p is ______________________

1. Set up H_0 and H_1:

   Assume H_0.

2. The significance level is:

3. Therefore the critical region is:

4. Since _________________, the result is significant / not significant, and the null hypothesis is rejected / not reject
   there is sufficient / insufficient evidence, at the ____ level, to say ______________________