Probability of Type II Errors

A type II error occurs if the true mean is anything other than what is suggested by the null hypothesis, and the null hypothesis has not been rejected. The probability of a type II error is dependent on the true value of the population parameter. If the true mean is very different to that in the null hypothesis, it would be highly likely that this would be detected in a hypothesis test. However, if the true mean is different but still very close to that in the null hypothesis, it would be much harder to detect - making the probability of a type II error being very high.

In a hypothesis test:

$$\text{P(type II error)} = \text{P(not rejecting } H_0 \text{|the true mean is a specified different value to } H_0 \text{)}$$

Example 1: The mass of eggs laid by one hen is known to have a normal distribution with variance 8 and an unknown mean. A hypothesis test is conducted to test the hypotheses $H_0: \mu = 40;\ H_1: \mu \neq 40$, using the mean of a sample size of 12 and 5% significance level.

a) Find the critical regions required for this test.
b) Given that the true mean is 38.2, calculate the probability of making a type II error.

a) Define the variables.	$X \sim N(\mu, 8)$ For the hypotheses $H_0: \mu = 40;\ H_1: \mu \neq 40$ $\bar{X} \sim N\left(40, \frac{8}{12}\right)$
This is a two tailed test. Use a calculator to find the values for the critical region.	$P(\bar{X} < a) = 0.025$, so $a = 38.400\ to\ 3dp.$ $P(\bar{X} < b) = 0.975$, so $b = 41.600\ to\ 3dp.$ We do not reject H_0 if $38.400 < \bar{X} < 41.600$
b) If the actual mean is 38.2, we need to calculate the probability that the test statistic will fall within the acceptance region, given that it has a different mean.	$\text{P(type II error)} = P(38.4 < \bar{X} < 41.6)$ when $\bar{X} \sim N\left(38.2, \frac{8}{12}\right)$ $= P(\bar{X} < 41.6) - P(\bar{X} < 38.4)$ $= 0.4032$

Example 2: It is proposed to conduct a hypothesis test to see whether a coin is fair or not. The coin is flipped ten times, and if there are less than 3 observations of 'heads', the probability of flipping a head is said to be less than half.

a) State the appropriate hypotheses for this test.
b) Calculate the significance level of the test.
c) If the true probability of getting heads is 0.4, calculate the probability of a type II error.

a) This is a one-tailed hypothesis test, and so the hypotheses are given by.	$H_0: p(\text{heads}) = 0.5;\ H_1: p(\text{heads}) < 0.5$
b) The null hypothesis is to be rejected if there are less than three observations of heads.	If X is the number of heads seen in ten coin flips, and if H_0 is true then $X \sim B(10, 0.5)$. For this distribution $P(X < 3) = P(X \leq 2) = 0.0546$ Therefore, the significance level of the test is 5.46%.
c) The probability of type II error is given by the probability that the test statistic incorrectly falls in the acceptance region.	$\text{P(type II error)} = \text{P(accepting } H_0 \text{|} H_0 \text{ is false)}$ $P(X \geq 3) = 1 - P(X \leq 2)$ for $X \sim B(10, 0.4)$ $= 1 - 0.1672 = 0.833\ to\ 3dp$

The Power of a Hypothesis Test

The power of a hypothesis test is the probability that H_0 is correctly rejected. This is the equivalent to the probability of not making a type II error, and so is defined by:

$$\text{Power of a test} = 1 - \text{P(type II error)}$$

Example 3: A hypothesis test for a random variable with distribution given by $X \sim Po(6)$ is conducted, where H_0 is rejected for $X \geq 10$ or $X < 3$. If the real value of $\lambda = 4$, find:

a) The probability of a type II error.
b) The power of the test.

a) The probability of a type II error is the probability that the test statistic falls in the acceptance region, even though it has a different mean. The actual distribution is $X \sim Po(4)$ so using this, calculate $P(3 \leq X \leq 9)$.	The actual distribution is $X \sim Po(4)$ so calculate $P(3 \leq X \leq 9) = 0.7538$
b) The power of a test is given as the probability of not making a type II error.	Power of the test: $= 1 - P(3 \leq X \leq 9)$ $= 1 - 0.7538 = 0.2462$

Example 4: The lengths of a certain species of beetle is known to have a standard deviation of 4mm. A scientist wants to test whether the beetles have a length of 100mm on average. The scientist does this by collecting samples from 10 different beetles from a nest and performing a hypothesis test to the 5% significance level.

a) State the null and alternative hypotheses for this test.
b) Find the critical region for this test.
c) What is the probability of making a type I error in this test?
d) Given that the average lengths of the beetles is actually 102mm, calculate the probability of making a type II error in this test.
e) Hence calculate the power of the test.

a) This is a two tailed hypothesis test, so the relevant hypotheses are as follows.	$H_0 \mu = 100mm;\ H_1: \mu \neq 100mm$
b) The means should be normally distributed with the standard deviation given. As this is a two tailed test, there will be two values to find, corresponding to the 2.5% cut-off. This can be done using your calculator or looking up the Z-values in a table.	$\bar{X} \sim N\left(1.0, \frac{0.04^2}{10}\right)$ $P(\bar{X} < a) = 0.025$, so $a = 97.52\ to\ 2dp.$ $P(\bar{X} < b) = 0.975$, so $b = 102.48\ to\ 2dp.$ We do not reject H_0 if $97.52 < \bar{X} < 102.48$
c) The probability of type I error is the actual significance level of the test. In this example, the data is continuous, so we have a critical region that corresponds with exactly with the 5% significance level of the test.	The actual significance level of the test = 0.05 Therefore, the probability of type I error is also 0.05.
d) The probability of type II error is the probability that the test statistic will fall in the acceptance region for the hypothesis test, given that it has a different distribution. This distribution has the same standard deviation but with a different mean.	The actual distribution of the mean is given by $\bar{X} \sim N\left(1.02, \frac{0.04^2}{10}\right)$ $\text{P(type II error)} = \text{P(accepting } H_0 \text{|} H_0 \text{ is false)}$ $P(97.52 < \bar{X} < 102.48) = 0.6476$
e) The power of the test is given by $1 - \text{P(type II error)}$.	$\text{Power} = 1 - \text{P(type II error)}$ $= 1 - 0.6476 = 0.3524$