Exercise Set 7 - Solution

1 Photodetector Efficiency

a) In this case, we are given the mean $\bar{X} = 1.3A/W$ of our sample and the true standard deviation $\sigma = 0.1\mu_0 = 0.13A/W$ of our process, so we use the **z-test**. Furthermore, we want to make sure our efficiency is higher than the reference mean $\mu_0 = 1.2A/W$, so we use a **one-sided z-test**.

b) It is always good practice to formulate our hypothesis. Here the hypothesis is that our efficiency is higher, so $H_1 : \mu > 1.2A/W$ and hence $H_0 : \mu \leq 1.2A/W$. We also choose a level of significance $\alpha = 0.01$.

We can calculate the z-value for our statistic : $Z(X) = \frac{\bar{X}-\mu_0}{\sigma}\sqrt{4} = 1.54$. From here there are two ways of solving the problem : either compare the z-values, or compare the probabilities.

- Comparing the z-values. We calculate the z-value corresponding to a confidence level $\alpha = 0.01$, $z_{0.99} = 2.33$. This creates an exclusion interval $K = [2.33, +\infty[$.
  Here $Z(X) = 1.54 \notin K$ so we cannot refute H_0.
- Comparing the probabilities. From the z-value of our statistic, we can deduce a confidence level. Using the z-table, we have $z_{0.9382} = 1.54$, meaning we have a confidence level $\alpha_0 \approx 0.94 < 0.99$ for our true mean to actually be greater than μ_0.

In both cases, we conclude that our photodetector do not meet the standard with sufficient confidence level.

c) Here we want to know with which confidence level our value of the mean would be on the edge of refuting H_0. If we already computed the probability in exercise **b)**, we can already answer that $\alpha_{\text{limit}} \approx 0.94$. In this case, the probability that our probability is actually worse than the reference value is $P = 1 - 0.94 = 6\%$.

d) Now we want to know how many times we should do the experiment to be able to reject H_0 (assuming the sample mean stays the same). This means setting the condition :

$$\frac{\bar{X}-\mu_0}{\sigma}\sqrt{N} \geq z_{0.99} \Rightarrow N \geq (\frac{\sigma}{\bar{X}-\mu_0} z_{0.99})^2$$

Doing so we get $N \geq 9.25$, so we need to do at least 10 experiments.

2 Novel diets for a healthy lifestyle

We do not know the true variance of the diets, so we have to estimate the standard deviation from the sample. This means we need to run two **one-sided T-test**. Our hypothesis is that they lost weight, such that $H_1 : \mu < 105kg$ and hence $H_0 : \mu \geq 105kg$.

First, we calculate the sample parameters for the 2 series, with s_i calculated with the unbiased standard deviation $s_{1,2}^2 = \frac{\sum(\bar{X}_{1,2}-x_i)^2}{N-1}$:

$$\bar{X}_1 = 93.0 \qquad s_1^2 = 48.18 \qquad\qquad \bar{X}_2 = 103.6 \qquad s_2^2 = 73.17$$

Second, we calculate the t-statistics:

$$t_1 = \frac{\bar{X}_1 - 105}{s_1}\sqrt{12} = -5.99 \qquad\qquad t_2 = \frac{\bar{X}_2 - 105}{s_2}\sqrt{12} = -0.55$$

Third, we compute the critical regions. As low values of t disprove H_0 in our one-sided test, we need to compute $t_{0.05,11}$, the t-value for $\alpha = 0.05$ with 11 degrees of freedom. We can look at the t-table and find $t_{0.05,11} = -1.796$.

Diet 1 clearly falls deep into the critical region, such that H_0 must be rejected. This appears to be a very effective diet. For fun, we could compute the p-value and find $p = 4.52 \cdot 10^{-5}$. The probability that this is compatible with H_0 is very small, and this makes sense when one looks at the data. Most weights are significantly below 105, only one at 105, and none above. Diet 2 still has an average below 105, but the scatter is so high that we cannot conclude this is due to the diet. The T-value is very low, deep in the acceptance region, thus it is very likely that this small deviation from 105 is only due to random fluctuations and that, in fact, our diet 2 had no effect.

3 Error of type 1 (false discovery/positive) and 2 (false negative), power of a test

a) We have to compute the probability that the statistic $\sqrt{n}\bar{Y}$ falls outside of $[-2, 2]$. This probability is the area under the bell curve of the centred of reduced normal law $\mathcal{N}(0, 1)$ between $-\infty$ and -2 plus the area between 2 and ∞. This simply is 1 minus the quantile associated with $z = 2$, times 2. So $2 \cdot (1 - \Phi(2)) = 2 - 2 \cdot 0.9772 = 4.56\%$.

b) When taking 9 measurements, the mean follows the distribution $\mathcal{N}(0, 1/9)$. 2 has to be reduced by being divided by the standard deviation. This implies that the probability of a) does not change. This simply reflects the fact that taking more measurements sharpens the distribution function in the same way as the window of acceptance of our test shrinks.

c) A type 2 error means that the test accepts the null hypothesis $\theta = 0$ while it is not true. Lets assume that it is not true, and the real $\theta = 1$. As we saw in b), the random variable $\bar{Y}$ then follows the distribution $\mathcal{N}(1, 1/9)$.

$$P(\text{Error type 2}) = P(|\sqrt{n} \cdot \bar{Y}| \leq 2) = P(|\bar{Y}| \leq \frac{2}{3})$$

To evaluate this, we need the CDF of the distribution function of $\bar{Y}$, $\mathcal{N}(1, 1/9)$. Let us denote this CDF by $\tilde{\Phi}$. Then,

$$P(|\bar{Y}| \leq \frac{2}{3}) = \tilde{\Phi}(\frac{2}{3}) - \tilde{\Phi}(-\frac{2}{3})$$

We first have to transform $\bar{Y}$ to the standard normal variable $Z = \frac{\bar{Y}-\theta}{\frac{1}{3}} = 3(\bar{Y} - \theta)$. For this variable, we can use the standard CDF Φ. We find:

$$P(|\bar{Y}| \leq \frac{2}{3}) = \Phi(2 - 3\theta) - \Phi(-2 - 3\theta)$$

The z-table (or calculator or computer) gives us 15.9% for $\theta = 1$.

Our null hypothesis is $H_0 : \theta = 0$. If the test rejects H_0 while in reality $\theta = 0$, the test does a type I error. If the test accepts H_0, thus claiming that $\theta = 0$ while it is in reality not, then it makes a type II error.

d) Using the same reasoning as before:

θ	Type 1	Type 2	Power
0	4.56%	-	-
0.5	-	69.1%	30.9%
1	-	15.9%	84.1%
1.5	-	0.62%	99.38%
2	-	$\sim$ 0.00%	100%

Table 1: Type 1 (false discovery/positive) and 2 (false negative) error probabilities and associated power.

The power is the probability the test rejects the null hypothesis ($\theta = 0$) when $\theta \neq 0$, so it is $1 - P(\text{Error type 2})$.

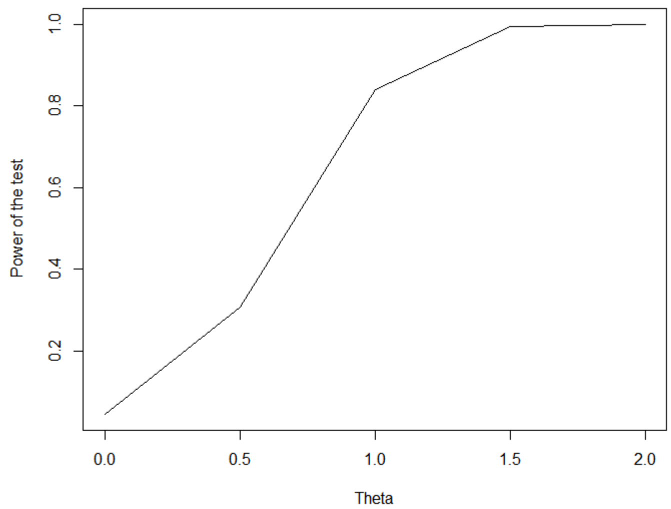

Figure 1: Power of the test for different θ.

In the graph, for $\theta = 0$, we plot the level of the test, which is the probability the test rejects the null hypothesis ($\theta = 0$) when $\theta = 0$. Note that all these numbers critically depend on the number of experiments, n.