Exercise Set 7 - Solution

1 Exam-style Python questions [normal-advanced]

a) The loop will return the difference between a point and the mean for each elements in the array :

0.0

1.0

-1.0

b) This will return 0.

c) Comparator operations are ill-defined for arrays (should they be applied to each element or to the whole array somehow?). To solve use np.vectorize(myFilter)(myArray) or make a loop (either inside or outside the function) over each element or use a logical_and type operation

2 Exam grades [basic]

Mean: 4.791 (you can round to 4.8)

Median: 5.0

Unbiased Variance: 0.61

Although the median is quite close to the mean (and indeed the distribution is somewhat symmetric around the median), there is a minimum in frequency close to the median. In a normal distribution, you would not expect this.

We will learn how to make such a distinction quantitatively, but by visual inspection it seems that the sum of two normal distributions, one centred around 4.3 and the other centred close to 5.3, would give a better description of the data. Such distributions are known as *bimodal* distributions.

3 Novel diets for a healthy lifestyle

Our hypothesis is that they lost weight, such that $H_1 : \mu < 105kg$ and hence the null hypothesis is $H_0 : \mu \geq 105kg$.

We do not know the true variance of the diets, so we have to estimate the variance from the sample. This means we need to run, for each diet separately, a **one-sided T-test**.

First, we calculate the sample parameters for the 2 series, with s_i calculated with the unbiased variance $s_{1,2}^2 = \frac{\sum(x_i - \bar{x}_{1,2})^2}{n-1}$ (as always it can be faster to sum x^2 itself and use the usual equality):

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

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.

4 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:

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

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.

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.

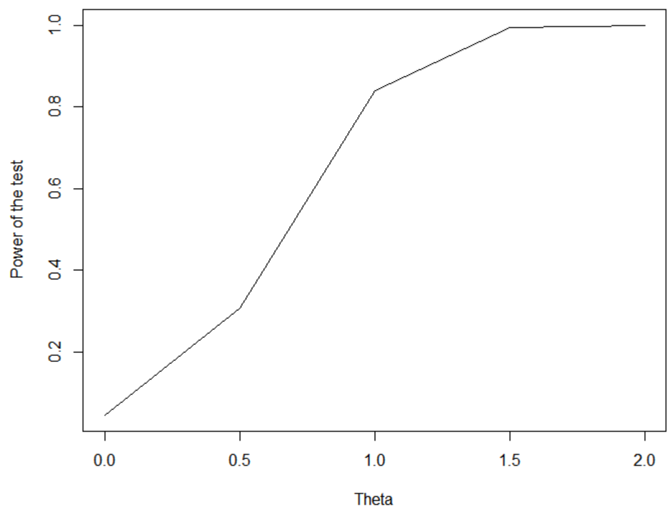

Figure 1: Power of the test for different θ.