# Data Science and Statistical Computing

# Tutorial 1, Week 2 Solutions

## Q1

### (a)

This is just a Binomial probability,

$$\binom{10}{3} \left(\frac{1}{10}\right)^3 \left(1 - \frac{1}{10}\right)^{10-3} = 0.0574$$

### (b)

5 samples are $\leq 1.48$, so

$$\sum_{k=0}^{2} \binom{10}{k} \left(\frac{5}{10}\right)^k \left(1 - \frac{5}{10}\right)^{10-k} = 0.0547$$

### (c)

5 samples are $\leq 1.48$, 2 samples are $> 1.52$, so

$$\binom{10}{2} \left(\frac{5}{10}\right)^2 \left(\frac{2}{10}\right)^{10-2} = 0.000029$$

## Q2

### (a)

- Yes, the suggestion is fine.
- A test statistic just has to be any function of the data which will be extreme when the null is wrong and not-extreme when the null is true.
- $\bar{X} = \lambda$, so
  - the sum will be extreme relative to $\lambda$ (i.e., larger than $\lambda$) when the $H_1$ holds
  - the sum will be close to $\lambda$ when $H_0$ holds
- A valid critique would be that this test statistic is not a pivotal quantity (i.e., the distribution of the test statistic is not independent of $\lambda$), but since we're Monte Carlo testing this does not concern us particularly.

### (b)

$$t_{\text{obs}} = \sum_{i=1}^{3} x_i = 2 + 0 + 1 = 3$$

**(c)**

$N = 10$, because Monte Carlo testing of a data set of size $n$ involves drawing $N$ samples of size $n$. Here, $n = 3$ and there are 30 simulated Poisson draws, so $N = 10$.

**(d)**

Taking the values grouped as they appear in the question output (where each row of output contains two simulated datasets of size $n = 3$), we have:

$$\mathbf{t} = (3, 2, 4, 4, 6, 4, 2, 4, 1, 3)$$

These are all iid samples, so any other grouping that does not depend on the sampled values is valid!

**(e)**

$$\begin{aligned}
\hat{p} &= \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\{t_i \geq t_{\text{obs}}\} \\
&= \frac{1}{10}\big(\mathbb{1}\{3 \geq 3\} + \mathbb{1}\{2 \geq 3\} + \mathbb{1}\{4 \geq 3\} + \cdots + \mathbb{1}\{3 \geq 3\}\big) \\
&= \frac{7}{10} = 0.7
\end{aligned}$$

**(f)**

$\hat{p} > \alpha$, so we do <u>not</u> reject the null hypothesis. There is no evidence to suggest that the parameter $\lambda$ exceeds 1.

## Q3

**(a)**

We need to compute
$$p = \mathbb{P}(T \geq t_{\text{obs}} \mid H_0 \text{ true})$$
where $T = \sum_{i=1}^{3} X_i$.

$H_0 : \lambda = 1$ and $n = 3$, so based on the hint we know the distribution of $T$ is Poisson with $\lambda = 3 \times 1 = 3$.

Therefore,

$$\begin{aligned}
p = \mathbb{P}(T \geq t_{\text{obs}} \mid H_0 \text{ true}) &= \mathbb{P}\big(T \geq 3 \mid T \sim \text{Pois}(\lambda = 3)\big) \\
&= 1 - \sum_{k=0}^{2} \frac{\lambda^k}{k!} e^{-\lambda} \\
&= 1 - (0.0498 + 0.1494 + 0.2240) \\
&= 0.5768
\end{aligned}$$

**(b)**

This is not a trick question, just testing understanding! This is precisely the definition of the p-value, so is 0.5768 as just calculated in (a)

**(c)**

This is now repeated trials of the Monte Carlo simulation and so is just Binomial($N, p = 0.5768$).

**(d)**

In Q2, we were testing at a significance level $\alpha = 0.1$ using just $N = 10$ Monte Carlo simulations of the test statistic.

Here, $p > \alpha$ and the correct conclusion of the test is to not reject the null. Therefore, the resampling risk in this situation is $\mathbb{P}(\hat{p} \leq \alpha)$.

Based on (c) we know the distribution of the number of Monte Carlo trials exceeding $t_{\text{obs}}$ is Binomial($N = 10, p = 0.5768$) and we know that

$$\hat{p} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{t_i \geq t_{\text{obs}}\}$$

Therefore,

$$\begin{aligned} \mathbb{P}(\hat{p} \leq \alpha) &= \mathbb{P}(\hat{p} \leq 0.1) \\ &= \mathbb{P}\left( \sum_{i=1}^{10} \mathbb{1}\{t_i \geq t_{\text{obs}}\} \leq 1 \right) \\ &= \sum_{k=0}^{1} \binom{10}{k} 0.5768^k (1 - 0.5768)^{10-k} \\ &= 0.0027 \end{aligned}$$

So there is a 0.3% chance that a Monte Carlo test on this problem will fail to give the correct conclusion. On one hand this is not bad given the ludicrously small sample, but on the other the true p-value is very far from the significance level. In practice, the resampling risk can often be much higher than this, so don't assume this is typical.

**Important:** this is the chance of "failing" in the sense of agreeing with the exact p-value (i.e., if we know the exact distribution of the test statistic). Of course, the exact test could still be making a Type II error!

## Q4

**(a)**

The lifetimes in weeks are, say, $a_i = \frac{x_i}{7}$ and $b_i = \frac{y_i}{7}$. Therefore,

$$\begin{aligned} \bar{a} &= \frac{1}{n} \sum a_i \\ &= \frac{1}{7n} \sum x_i \\ &= \frac{\bar{x}}{7} \end{aligned}$$

This sample mean has standard error,

$$
\begin{aligned}
se_{\bar{a}} = \frac{s_a}{\sqrt{n}} &= \frac{1}{\sqrt{n}}\sqrt{\frac{1}{n-1}\sum(a_i - \bar{a})^2} \\
&= \frac{1}{\sqrt{n}}\sqrt{\frac{1}{7^2(n-1)}\sum(x_i - \bar{x})^2} \\
&= \frac{s_x}{7\sqrt{n}} = \frac{se_{\bar{x}}}{7}
\end{aligned}
$$

Similarly for $b_i, y_i$. Therefore, both sample mean and its standard error are simply divided by 7.

**(b)**

The standard error of the difference is,

$$
\begin{aligned}
se_{\bar{a}-\bar{b}} &= \sqrt{se_{\bar{a}}^2 + se_{\bar{b}}^2} \\
&= \frac{\sqrt{se_{\bar{x}}^2 + se_{\bar{y}}^2}}{7} \\
&= \frac{se_{\bar{x}-\bar{y}}}{7}
\end{aligned}
$$

Thus,

$$
\frac{\bar{a} - \bar{b}}{se_{\bar{a}-\bar{b}}} \equiv \frac{\bar{x} - \bar{y}}{se_{\bar{x}-\bar{y}}}
$$

In other words, as expected this is scale invariant.

**(c)**

Let the new dataset with repeats be $\{z_i : i \in \{1, \ldots, nN\}\}$ where each $x_i$ is repeated $N$ times.

$$
\begin{aligned}
\bar{z} &= \frac{1}{nN}\sum_{i=1}^{nN} z_i \\
&= \frac{1}{nN}\sum_{i=1}^{n} N x_i \\
&= \bar{x}
\end{aligned}
$$

$$
\begin{aligned}
se_{\bar{z}} = \frac{s_z}{\sqrt{nN}} &= \frac{1}{\sqrt{nN}}\sqrt{\frac{1}{nN-1}\sum_{i=1}^{nN}(z_i-\bar{z})^2} \\
&= \frac{1}{\sqrt{nN}}\sqrt{\frac{N}{nN-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} \\
&= \frac{1}{\sqrt{nN}}\sqrt{\frac{nN-N}{nN-1}\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} \\
&= \frac{1}{\sqrt{nN}}\sqrt{\frac{nN-N}{nN-1}}s_x \\
&= \frac{s_x}{\sqrt{n}}\sqrt{\frac{n-1}{nN-1}} = se_{\bar{x}}\sqrt{\frac{n-1}{nN-1}}
\end{aligned}
$$

Yes, $\sqrt{\frac{n-1}{nN-1}}$ behaves roughly like $N^{-1/2}$.