# Assignment 7 for Statistical Computing and Empirical Methods: Sample and population

Henry Reeve

## Introduction

This document describes your seventh assignment for Statistical Computing and Empirical Methods (Unit EMATM0061) on the MSc in Data Science. Before starting the assignment it is recommended that you first watch video lecture 7 entitled “Sample and population”.

It is recommended that you write your answers within an RMarkdown file using TEX https://en.wikipedia.org/wiki/TeX (https://en.wikipedia.org/wiki/TeX). However, this is optional.

## 1 Basic properties of random variables

Recall from the lectures that a random variable $X$ with outcome space $\mathcal{S}$ satisfies the following basic properties:

1. $\mathbb{P}(X \in \mathcal{S}) = 1$,
2. Given all well-behaved sets $A \subseteq \mathcal{S}$, we have $\mathbb{P}(X \in A) \in [0,1]$,
3. Given sets $A_1, A_2, \cdots, A_k \subseteq \mathcal{S}$ with $A_i \cap A_j = \emptyset$ for $i \neq j$ we have,

$$\mathbb{P}\left(X \in A_1 \cup \cdots \cup A_k\right) = \mathbb{P}\left(X \in A_1\right) + \cdots + \mathbb{P}\left(X \in A_k\right).$$

Recall that $\emptyset$ denotes the empty set, the unique set which contains no points. Apply the basic properties of random variables to show that $\mathbb{P}(X \in \emptyset) = 0$.

**A** Let $A_1 = \emptyset$ and $A_2 = \mathcal{S}$. Note that $A_1 \cup A_2 = \mathcal{S}$ and $A_1 \cap A_2 = \emptyset$. Hence, by properties 1 and 3 we have

$$\begin{aligned}
1 &= \mathbb{P}(X \in \mathcal{S}) = \mathbb{P}(X \in A_1 \cup A_2) \\
&= \mathbb{P}(X \in A_1) + \mathbb{P}(X \in A_2) = \mathbb{P}(X \in \emptyset) + \mathbb{P}(X \in \mathcal{S}) \\
&= \mathbb{P}(X \in \emptyset) + 1.
\end{aligned}$$

Subtracting $1$ from both sides gives $\mathbb{P}(X \in \emptyset) = 0$.

Now, suppose that $A \subseteq \mathcal{S}$ and $B \subseteq \mathcal{S}$ are subsets of $\mathcal{S}$. What can we say about $\mathbb{P}\left(X \in A \cup B\right)$ in terms of $\mathbb{P}\left(X \in A\right)$ and $\mathbb{P}\left(X \in B\right)$?

**A** First note that we don't always have $\mathbb{P}\left(X \in A \cup B\right) = \mathbb{P}(X \in A) + \mathbb{P}(X \in B)$ and property 3 requires that $A \cap B = \emptyset$. Consider for example that $A = B = \mathcal{S}$. We do always have the following inequalities

$$\max\{\mathbb{P}(X \in A), \mathbb{P}(X \in B)\} \leq \mathbb{P}(X \in A \cup B) \leq \mathbb{P}(X \in A) + \mathbb{P}(X \in B)$$

Both inequalities from the fact proved in the lecture that $C \subset D \subset \mathcal{S}$ implies $\mathbb{P}(X \in C) \leq \mathbb{P}(X \in D)$. Indeed, to prove the first inequality we first apply this inequality with either $D = A \cup B$ and $C = A$ or $C = B$. To prove the second inequality we note that $B \backslash A = \{s \in B : s \notin A\} \subseteq B$ and $(B \backslash A) \cap A = \emptyset$. Thus by property 3 we have

$$\begin{aligned}\mathbb{P}(X \in A \cup B) &= \mathbb{P}\left(X \in A \cup (B \backslash A)\right) \\ &= \mathbb{P}\left(X \in A\right) + \mathbb{P}\left(X \in B \backslash A\right) \\ &\leq \mathbb{P}\left(X \in A\right) + \mathbb{P}\left(X \in B\right),\end{aligned}$$

where the final inequality applies the above fact with $C = B \backslash A$ and $D = B$.

## 2 Expectation and variance

Suppose that $\alpha$,$\beta \in [0, 1]$ with $\alpha + \beta \leq 1$. Suppose that $X$ is a random variable with outcome space $\mathcal{S} = \{0, 1, 5\}$. Suppose that $\mathbb{P}\left(X = 1\right) = \alpha$ and $\mathbb{P}\left(X = 5\right) = \beta$.

What is the probability mass function $p : \mathcal{S} \rightarrow [0, 1]$ for $X$?

**A** We have

$$p(x) = \begin{cases} 1 - \alpha - \beta & \text{if } x = 0 \\ \alpha & \text{if } x = 1 \\ \beta & \text{if } x = 5 \\ 0 & \text{otherwise.} \end{cases}$$

What is the expectation of $X$?

**A**

$$\mathbb{E}[X] = \alpha + 5\beta.$$

What is the variance of $X$?

**A**

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[X^2] - E[X]^2 \\ &= (\alpha + 25\beta) - (\alpha + 5\beta)^2 \\ &= \alpha + 25\beta - \alpha^2 - 25\beta^2 - 10\alpha\beta.\end{aligned}$$

## 3 Transformations of random variables

Suppose that $X$ is a discrete random variable with outcome space $\mathcal{S}_X$ and probability mass function $p_X : \mathcal{S}_X \to [0,1]$. Suppose that $S_Y$ is another space and $\varphi : \mathcal{S}_X \to \mathcal{S}_Y$ is a function. For each $A \subseteq \mathcal{S}_Y$ we define $\varphi^{-1}(A) \subseteq \mathcal{S}_X$ by $\varphi^{-1}(A) := \{x \in \mathcal{S}_X : \varphi(x) \in A\}$.

We define a new random variable $Y = \varphi(X)$ by $\mathbb{P}\left(\varphi(X) \in A\right) = \mathbb{P}\left(X \in \varphi^{-1}(A)\right)$ for $A \subseteq \mathcal{S}_Y$.

Prove that $Y = \varphi(X)$ satisfies the three basic properties of a random variable with outcome space $\mathcal{S}_Y$.

### A

1. We have $\varphi^{-1}(\mathcal{S}_Y) = \{x \in \mathcal{S}_X : \varphi(x) \in \mathcal{S}_Y\} = \mathcal{S}_X$, so by property 1 for $X$,

$$\mathbb{P}[Y \in \mathcal{S}_Y] = \mathbb{P}[\varphi(X) \in \mathcal{S}_Y] = \mathbb{P}[X \in \varphi^{-1}(\mathcal{S}_Y)] = \mathbb{P}[X \in \mathcal{S}_X] = 1.$$

2. Given any $A \subseteq \mathcal{S}_Y$ we have $\varphi^{-1}(A) \subseteq \mathcal{S}_X$, so by property 2 for $X$ we have, $\mathbb{P}[Y \in A] = \mathbb{P}[X \in \varphi^{-1}(A)] \in [0,1]$.
3. Suppose $A_1, \cdots, A_k \subseteq \mathcal{S}_Y$ with $A_i \cap A_j = \emptyset$ for $i \neq j$. Then $\varphi^{-1}(A_1), \cdots, \varphi^{-1}(A_k) \subseteq \mathcal{S}_Y$ and for $i \neq j$, $\varphi^{-1}(A_i) \cap \varphi^{-1}(A_j) = \emptyset$ (if $x \in \varphi^{-1}(A_i) \cap \varphi^{-1}(A_j)$ then $\varphi(x) \in A_i \cap A_j$). Moreover, $x \in \varphi^{-1}\left(A_1 \cup \cdots A_k\right)$ if and only if $\varphi(x) \in A_1 \cup \cdots A_k$ which holds if and only if $\varphi(x) \in A_i$ for some $i = 1, \cdots, k$, which in turn holds if and only if $x \in \varphi^{-1}(A_1) \cup \cdots \cup \varphi^{-1}(A_k)$. Thus, $\varphi^{-1}\left(A_1 \cup \cdots A_k\right) = \varphi^{-1}(A_1) \cup \cdots \cup \varphi^{-1}(A_k)$. Hence, by property 3 for $X$ we have,

$$\begin{aligned}\mathbb{P}\left[Y \in A_1 \cup \cdots \cup A_k\right] &= \mathbb{P}\left[\varphi(X) \in A_1 \cup \cdots \cup A_k\right] \\ &= \mathbb{P}\left[X \in \varphi^{-1}\left(A_1 \cup \cdots \cup A_k\right)\right] \\ &= \mathbb{P}\left[X \in \varphi^{-1}\left(A_1\right) \cup \cdots \cup \varphi^{-1}\left(A_k\right)\right] \\ &= \mathbb{P}\left[X \in \varphi^{-1}\left(A_1\right)\right] + \cdots + \mathbb{P}\left[X \in \varphi^{-1}\left(A_k\right)\right] \\ &= \mathbb{P}\left[\varphi(X) \in A_1\right] + \cdots + \mathbb{P}\left[\varphi(X) \in A_k\right] \\ &= \mathbb{P}\left[Y \in A_1\right] + \cdots + \mathbb{P}\left[Y \in A_k\right].\end{aligned}$$

Write the probability density function $p_Y : \mathcal{S}_Y \to [0,1]$ of $Y = \varphi(X)$ in terms of $p_X$.

**A**

For each $u \in \mathcal{S}_Y$,

$$\begin{aligned} p_Y(u) &= \mathbb{P}[Y = u] = \mathbb{P}[\varphi(X) = u] \\ &= \mathbb{P}[\varphi(X) \in \{u\}] = \mathbb{P}[X \in \varphi^{-1}(\{u\})] \\ &= \sum_{v \in \varphi^{-1}(\{u\})} p_X(v) = \sum_{\varphi(v)=u} p_X(v). \end{aligned}$$

## 4 Poisson random variables

Many discrete random variables have a finite outcome space. Poisson random variables are a family of discrete real-valued random variables with outcome space $\mathbb{N}_0 := \{0, 1, 2, 3, \cdots\}$. Take $\lambda > 0$. The Poisson random variable $X$ with parameter $\lambda$ has probability mass function $p_\lambda : \mathbb{N}_0 \to (0, \infty)$ defined by

$$p_\lambda(k) = \frac{\lambda^k e^{-\lambda}}{k!} \text{ for all } k \in \mathbb{N}_0.$$

Show that $p_\lambda$ is a well-defined probability mass function.

**A** We note that $p_\lambda(k) \geq 0$ for all $k \in \mathbb{N}$ and

$$\sum_{k \in \mathbb{N}_0} p_\lambda(k) = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \left( \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right) = e^{-\lambda} \cdot e^{\lambda} = 1,$$

where we use the power series for the exponential function.

Compute both the expectation and the variance of a Poisson random variable $X$ with probability mass function $p_\lambda$.

**A**

We can compute the expectation,

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k=0}^{\infty} k \cdot p_\lambda(k) = \sum_{k=1}^{\infty} k \cdot p_\lambda(k) \\ &= \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-1)!} \\ &= \lambda \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1} e^{-\lambda}}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda.\end{aligned}$$

In addition we have

$$\begin{aligned}\mathbb{E}[X^2] &= \sum_{k=0}^{\infty} k^2 \cdot p_\lambda(k) = \sum_{k=1}^{\infty} k^2 \cdot p_\lambda(k) \\ &= \sum_{k=1}^{\infty} k^2 \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{(k-1)!} \\ &= \lambda \cdot \sum_{j=0}^{\infty} (j+1) \frac{\lambda^j e^{-\lambda}}{j!} \\ &= \lambda \cdot \sum_{j=0}^{\infty} (j+1) \cdot p_\lambda(j) \\ &= \lambda \cdot \left\{ \left( \sum_{j=0}^{\infty} j \cdot p_\lambda(j) \right) + \left( \sum_{j=0}^{\infty} p_\lambda(j) \right) \right\} \\ &= \lambda \cdot (\mathbb{E}[X] + 1) = \lambda(\lambda + 1).\end{aligned}$$

Thus, $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda$.