Uniform distributions

David J. H. Shih

Intended learning outcomes

- Recognize and apply uniform distributions in statistical models
- Prove that Kolmogorov's axioms hold (or not) for given functions
- Derive expected values of random variables
- Apply empirical probability mass functions in models

Discrete vs. continuous random variables

Definitions

A random variable X is *continuous* if its cdf $F_X(x)$ is a continuous function of x.

A random variable X is *discrete* if $F_X(x)$ is a step funtion of x.

Discrete uniform distribution

Definition

A random variable X has a discrete uniform distribution if

$$P\left(X = x \mid N\right) \; = \; \frac{1}{N}, \quad x \in \{1, 2, \dots, N\},$$

where $N > 0$ is an integer parameter.
Useful as an expression of ignorance.
Other possible notations are

$$\begin{aligned} p\left(X = x \mid N\right) &= \mathsf{Uniform}\left(x \mid N\right) \\ X &\sim \mathsf{Uniform}\left(N\right) \end{aligned}$$

Examples

- outcome of a fair coin flip
- outcome of a fair dice roll
- number on a drawn card from a well-shuffled deck
- biological sex of humans

Kolmogorov's axioms of probability

Given a sample space $\mathcal{S}$, a probability function P satisfies

1. $P(\mathcal{A}) \geq 0 \quad \forall \mathcal{A} \subseteq S.$ (non-negativity)
2. $P(\mathcal{S}) = 1.$ (unit measure)
3. For any $\mathcal{A}_1, \mathcal{A}_2, ... \subseteq S$
   s.t. $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset \; \forall i \neq j$,
   $$P\left(\bigcup_{i=1}^{\infty} \mathcal{A}_i\right) = \sum_{i=1}^{\infty} P(\mathcal{A}_i).$$ (additivity)

Remark: Given the sample space $\mathcal{X}$ of the random variable X, the induced probability function P_X must be defined properly so that it satisifies Kolmogorov's axioms.

Proof: Uniform$(x \mid N)$ is a probability distribution

Prove that $P_X(x) = \text{Uniform}(N)$ satisifies Kolmogorov's axioms of probability.

Non-negativity

For any $\mathcal{A} \subseteq \mathcal{X}$ where $A = \{a_1, \ldots a_K\}$, $P_X(\mathcal{A}) = \sum_k P_X(a_k)$ by the additivity axiom.
Further, $\sum_k P_X(a_k) = \sum_{k=1}^{K} \frac{1}{N} = \frac{K}{N}$.
Finally, $\frac{K}{N} > 0$ since $N > 0$ and $K > 0$.
Therefore, $P_X(\mathcal{A}) \geq 0$ for any $\mathcal{A} \subseteq \mathcal{X}$.

Proof: Uniform$(x \mid N)$ is a probability distribution

Unit measure

Given $\mathcal{X} = \{x_1, \ldots, x_N\}$,

$$\begin{aligned} P_X(\mathcal{X}) &= P_X(\{x_1, \ldots, x_N\}) \\ &= \sum_{i=1}^{N} P_X(x_i) \quad \text{(additivity)} \\ &= \sum_{i=1}^{N} \frac{1}{N} = \frac{N}{N} = 1 \end{aligned}$$

Proof: Uniform$(x \mid N)$ is a probability distribution

Recall that, for a discrete random variable X,

$$P_X(x) = P\left(\{s \in \mathcal{S} : X(s) = x\}\right).$$

It is a given that $P(s)$ is a probability function.

Additivity

Consider any $\mathcal{A}, \mathcal{B} \subseteq \mathcal{X}$ s.t. $\mathcal{A} \cap \mathcal{B} = \emptyset$.
Note that $\mathcal{A} = a_1, \ldots, a_K$ and $\mathcal{B} = b_1, \ldots, b_L$.
Define $c_1 = a_1, \ldots, c_K = a_K, c_{K+1} = b_1, \ldots, c_{K+L} = b_L$.

Proof outline

We will show that $P_X(\mathcal{A}) = \frac{K}{N}$, and similarly, $P_X(\mathcal{B}) = \frac{L}{N}$.
We will then show $P_X(\mathcal{A} \cup \mathcal{B}) = \frac{K}{N} + \frac{L}{N}$.
Therefore, $P_X(\mathcal{A} \cup \mathcal{B}) = P_X(\mathcal{A}) + P_X(\mathcal{B})$.
Our derivation can be extended to any number of subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ of $\mathcal{X}$, which completes the proof.

$$
\begin{aligned}
P_X(\mathcal{A}) &= P_X\left(\bigcup_{k=1}^{K} a_k\right) \\
&= P\left(\bigcup_{k=1}^{K} \{s_k \in \mathcal{S} : X(s_k) = a_k\}\right) \quad \text{(definition of } P_X\text{)} \\
&= \sum_{k=1}^{K} P\left(\{s_k \in \mathcal{S} : X(s_k) = a_k\}\right) \quad \text{(additivity of } P\text{)} \\
&= \sum_{k=1}^{K} P_X(a_k) \quad \text{(definition of } P_X\text{)} \\
&= \sum_{k=1}^{K} \frac{1}{N} \quad \text{(definition of Uniform}(x \mid N)\text{)} \\
&= \frac{K}{N}
\end{aligned}
$$

Similarly as above, we can also show $P_X(\mathcal{B}) = \frac{L}{N}$.

$$
\begin{aligned}
P_X\left(\mathcal{A} \cup \mathcal{B}\right) &= P_X\left(\left(\bigcup_{k=1}^{K} a_k\right) \cup \left(\bigcup_{l=1}^{L} b_l\right)\right) \qquad = \\
&= P\left(\bigcup_{k=1}^{K+L} \{s_k \in \mathcal{S} : X(s_k) = c_k\}\right) \quad \text{(definition of } c_k\text{)} \\
&= \sum_{k=1}^{K+L} P\left(\{s_k \in \mathcal{S} : X(s_k) = c_k\}\right) \quad \text{(additivity of } P\text{)} \\
&= \sum_{k=1}^{K+L} P_X(c_k) \quad \text{(definition of } P_X\text{)} \\
&= \sum_{k=1}^{K+L} \frac{1}{N} \quad \text{(definition of Discrete}(x \mid N)\text{)} \\
&= \frac{K}{N} + \frac{L}{N} = P_X(\mathcal{A}) + P_X(\mathcal{B}).
\end{aligned}
$$

Therefore, for any $\mathcal{A}, \mathcal{B} \subseteq \mathcal{X}$ s.t. $\mathcal{A} \cap \mathcal{B} = \emptyset$,

$$P_X\left(\mathcal{A} \cup \mathcal{B}\right) = P_X\left(\mathcal{A}\right) + P_X\left(\mathcal{A}\right).$$

Using similar steps as above, we can show for any pairwise disjoint subsets $\mathcal{A}_1, \mathcal{A}_2, ...$ of $\mathcal{X}$ that

$$P_X\left(\bigcup_{i=1}^{\infty} \mathcal{A}\right) = \sum_{i=1}^{\infty} P_X\left(\mathcal{A}_i\right). \quad \blacksquare$$

Further, using almost the same derivation as above, given that the probability function P on sample space $\mathcal{S}$ satisfies the addivity axiom, we can show that the previously defined $P_X(x)$ also satisfies the additivity axiom.

Probability distribution

Since the additvity axiom holds for any $P_X(x)$, to show that a given $P_X(x)$ correspond to a *proper* probability distribution, we just need to show that $P_X(x)$ satisfies the first two Komomgorov's axioms. Specifically,

1. $P_X(x) \geq 0 \quad \forall x \in \mathcal{X}.$

2. $\displaystyle\sum_{x \in \mathcal{X}} P_X(x) = 1$ if X is discrete;

   $\displaystyle\int_{\mathcal{X}} f_X(x)\, dx = 1$ if X is continuous.

Continuous uniform distribution

Definition

A random variable has a continuous uniform distribution over an interval $[a, b]$ if

$$f_X(x \mid a, b) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{otherwise,} \end{cases}$$

where $b > a$.
We can also write

$$\begin{aligned} p\left(X = x \mid a, b\right) &= \text{Uniform}\left(x \mid a, b\right) \\ X &\sim \text{Uniform}\left(a, b\right) \end{aligned}$$

Proof: Uniform(a, b) is probability distribution

Prove that $f_X(x) = \text{Uniform}(x \mid a, b)$ is a probability distribution

Non-negativity

If $x \notin [a, b]$, $f_X(x) = 0$.
If $x \in [a, b]$, $f_X(a) = \frac{1}{b-a} \geq 0$ since $b > a$.

Unit measure

$$\begin{aligned}
\int_{-\infty}^{\infty} f_X(x)\,dx &= \int_a^b f_X(x)\,dx = \int_a^b \frac{1}{b-a}\,dx \\
&= \left[\frac{x}{b-a}\right]_{x=b} - \left[\frac{x}{b-a}\right]_{x=a} \\
&= \frac{b}{b-a} - \frac{a}{b-a} = 1
\end{aligned}$$

Continuous uniform distribution

If model parameters a and b are spread further apart (at the same rate) to cover a larger interval $[a, b]$, we have

$$f_X(x \mid -a, a) = \frac{1}{2a},$$

where $b = -a$.

Since $\lim_{a \to \infty} f_X(x \mid -a, a) = 0$, a uniform distribution spanning $(-\infty, \infty)$ is *not* a proper probability distribution, due to violation of the unit measure axiom.

Instead, we sometimes define an improper uniform distribution by

$$p_X(x) \propto 1,$$

where is also not a proper distribution.

Expectation

Definition

The *expected value* or *mean* of a random variable $g(X)$ is

$$\mathsf{E}\left(g(X)\right) \triangleq \begin{cases} \int_{-\infty}^{\infty} g(x) f_X(x) dx & \text{if } X \text{ is continuous} \\ \sum_{x \in \mathcal{X}} g(x) f_X(x) & \text{if } X \text{ is discrete,} \end{cases}$$

provided that the integral or sum exists.
If $\mathsf{E}\left(g(X)\right) = \infty$, we say that $\mathsf{E}\left(g(X)\right)$ does not exist.

Example

Given $X \sim \mathsf{Uniform}(a, b)$, find $\mathsf{E}\left(X^2\right)$.

Variance

Definition

The *variance* of a random variable $g(X)$ is

$$\mathsf{Var}\left(X\right) \triangleq \mathsf{E}\left[\left(X - \mathsf{E}\left(X\right)\right)^2\right].$$

Alternative formula

Variance can be expressed as

$$\mathsf{Var}\left(X\right) = \mathsf{E}\left(X^2\right) - \left(\mathsf{E}(X)\right)^2.$$

This formula is useful in derivations, but calculating variance with this formula will lead to numerical instability.

Empirical probability mass function

Uniform distributions are convenient starting points if we have no prior knowledge about the random variable.

We should always ask if a distribution in a model is appropriate for the problem.

We can also estimate the pmf from the observed data.

Given *independent and identically distributed* (iid) samples $\{x_1, \ldots, x_N\}$, an **empirical pmf** for random variable X is given by

$$\hat{f}_N(x) = \frac{1}{N} \sum_{i=1}^{N} I(x_i = x),$$

where I is the indicator function.

Empirical cumulative distribution function

Given iid samples $\{x_1, \dots, x_N\}$, an **empirical cdf** for random variable X is given by

$$\hat{F}_N(x) = \frac{1}{N} \sum_{i=1}^{N} I(x_i \leq x).$$

The Glivenko-Cantelli theorem states that $\hat{F}_N(x)$ converges to $F(x)$. That is, as $N \to \infty$, with probability 1,

$$\sup_{x \in \mathcal{X}} \left| \hat{F}_N(x) - F(x) \right| = 0,$$

where $\sup$ is the supremum.

Summary

"Probability is a measure of our ignorance." - Richard Jeffreys

Casella & Berger 2002, sections 3.2, pages 85-86, 98-99.

Intended learning outcomes

- Prove that Kolmogorov's axioms hold (or not) for given functions
- Recognize and apply uniform distributions in statistical models
- Derive expected values of random variables
- Apply empirical probability mass functions in models