Probability: Axioms, Properties, Interpretations

Engineering Statistics
Section 2.2

Josh Engwer

TTU

03 February 2016

Chains of Unions, Chains of Intersections

Just as a **chain of sums** can be written compactly...

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_{n-1} + a_n \qquad \sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \cdots$$

...a **chain of unions** can similarly be compactly written...

$$\bigcup_{k=1}^{n} A_k = A_1 \cup A_2 \cup \cdots \cup A_{n-1} \cup A_n \qquad \bigcup_{k=1}^{\infty} A_k = A_1 \cup A_2 \cup A_3 \cup \cdots$$

...as can a **chain of intersections**...

$$\bigcap_{k=1}^{n} A_k = A_1 \cap A_2 \cap \cdots \cap A_{n-1} \cap A_n \qquad \bigcap_{k=1}^{\infty} A_k = A_1 \cap A_2 \cap A_3 \cap \cdots$$

...as can a **chain of products**.

$$\prod_{k=1}^{n} a_k = a_1 a_2 \cdots a_{n-1} a_n \qquad \prod_{k=1}^{\infty} a_k = a_1 a_2 a_3 \cdots$$

Probability Axioms

The **probability** of an event E, denoted $\mathbb{P}(E)$, is a real number.
It turns out only 3 **axioms** are needed to ensure probability behaves properly:

Proposition

(Probability Axioms)

Let $E \subseteq \Omega$ be an event from sample space Ω of an experiment.
Let $E_1, E_2, E_3, \ldots$ be an infinite collection of pairwise disjoint events. Then:

$(A1)$	$\mathbb{P}(E) \geq 0$	*Probablity of an Event is Non-Negative*
$(A2)$	$\mathbb{P}(\Omega) = 1$	*Probability of Sample Space is always One*
$(A3)$	$\mathbb{P}\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(E_k)$	*Probabilty of Infinite Disjoint Union is a Sum*

Basic Properties of Probability

The 3 probability axioms can be used to build more useful properties:

Proposition

(Basic Probability Properties)

Let $E \subseteq \Omega$ be an event from sample space Ω of an experiment.
Let $E_1, E_2, E_3, \ldots$ be an infinite collection of pairwise disjoint events. Then:

$(P1)$ $\mathbb{P}(\emptyset) = 0$ *Probability of Empty Set is always Zero*

$(P2)$ $\mathbb{P}\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mathbb{P}(E_k)$ *Probabilty of Finite Disjoint Union is a Sum*

Basic Properties of Probability

The 3 probability axioms can be used to build more useful properties:

Proposition

(Basic Probability Properties)

Let $E \subseteq \Omega$ be an event from sample space Ω of an experiment.
Let $E_1, E_2, \ldots, E_n$ be a finite collection of pairwise disjoint events. Then:

$(P1)$ $\quad \mathbb{P}(\emptyset) = 0$ *Probability of Empty Set is always Zero*

$(P2)$ $\quad \mathbb{P}\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mathbb{P}(E_k)$ *Probabilty of Finite Disjoint Union is a Sum*

PROOF: $(P1)$ Let $E_1 = E_2 = E_3 = \cdots = \emptyset$. Then $\bigcup_{k=1}^{\infty} E_k = \emptyset$
Moreover, the events are pairwise disjoint since $E_i \cap E_j = \emptyset \cap \emptyset \overset{S8}{=} \emptyset$ for $i \neq j$
$\therefore\ \mathbb{P}\left(\bigcup_{k=1}^{\infty} E_k\right) = \mathbb{P}(\emptyset) \overset{A3}{\Longrightarrow} \sum_{k=1}^{\infty} \mathbb{P}(E_k) = \mathbb{P}(\emptyset) \implies \sum_{k=1}^{\infty} \mathbb{P}(\emptyset) = \mathbb{P}(\emptyset)$
$\implies \mathbb{P}(\emptyset) = 0$ (as it's the only way for LHS & RHS to equal each other) $\square$

Basic Properties of Probability

The 3 probability axioms can be used to build more useful properties:

Proposition

(Basic Probability Properties)

Let $E \subseteq \Omega$ be an event from sample space Ω of an experiment.
Let $E_1, E_2, E_3, \ldots$ be an infinite collection of pairwise disjoint events. Then:

$(P1)$ $\mathbb{P}(\emptyset) = 0$ *Probability of Empty Set is always Zero*

$(P2)$ $\mathbb{P}\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mathbb{P}(E_k)$ *Probabilty of Finite Disjoint Union is a Sum*

PROOF: $(P2)$ Let $E_{n+1} = E_{n+2} = E_{n+3} = \cdots = \emptyset$

Then $\bigcup_{k=1}^{\infty} E_k = \left(\bigcup_{k=1}^{n} E_k\right) \cup \left(\bigcup_{k=n+1}^{\infty} \emptyset\right) \stackrel{S8}{=} \left(\bigcup_{k=1}^{n} E_k\right) \cup \emptyset \stackrel{S8}{=} \bigcup_{k=1}^{n} E_k$

Finally, apply (A3) and then (P1) □

More Properties of Probability

Proposition

(More Probability Properties)

Let $E \subseteq \Omega$ *be an event from sample space* Ω *of an experiment. Then:*

$$\begin{array}{lll} (P3) & \mathbb{P}(E^c) = 1 - \mathbb{P}(E) & \textit{Probability of Complement} \\ (P4) & \mathbb{P}(E) \leq 1 & \textit{Probability is never greater than One} \end{array}$$

More Properties of Probability

Proposition

(More Probability Properties)

Let $E \subseteq \Omega$ be an event from sample space Ω of an experiment. Then:

$$\begin{array}{lll} (P3) & \mathbb{P}(E^c) = 1 - \mathbb{P}(E) & \textit{Probability of Complement} \\ (P4) & \mathbb{P}(E) \leq 1 & \textit{Probability is never greater than One} \end{array}$$

<u>PROOF:</u>

$(P3)$ Note that $E \cup E^c = \Omega$ and $E \cap E^c = \emptyset \implies E, E^c$ are disjoint events.

Then $\mathbb{P}(E \cup E^c) = \mathbb{P}(\Omega) \overset{P2}{\Longrightarrow} \mathbb{P}(E) + \mathbb{P}(E^c) = \mathbb{P}(\Omega) \overset{A2}{\Longrightarrow} \mathbb{P}(E) + \mathbb{P}(E^c) = 1$

$\therefore \; \mathbb{P}(E^c) = 1 - \mathbb{P}(E) \qquad \square$

More Properties of Probability

Proposition

(More Probability Properties)

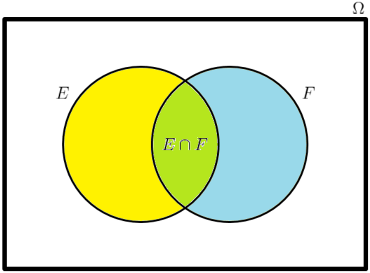
Let $E \subseteq \Omega$ *be an event from sample space* Ω *of an experiment. Then:*

$$\begin{array}{lll} (P3) & \mathbb{P}(E^c) = 1 - \mathbb{P}(E) & \textit{Probability of Complement} \\ (P4) & \mathbb{P}(E) \leq 1 & \textit{Probability is never greater than One} \end{array}$$

PROOF:

$(P4)$ By $(P3)$, $\mathbb{P}(E) + \mathbb{P}(E^c) = 1 \overset{A1}{\Longrightarrow} \mathbb{P}(E) \leq 1$ □

Probability of a Union of Two Events



Not Mutually Exclusive

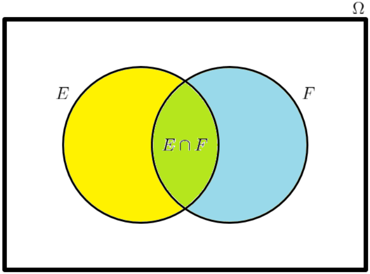
Proposition

(Probability of a Union of Two Events)

Let $E, F \subseteq \Omega$ be two events from sample space Ω of an experiment. Then:

$$(P5) \quad \mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F) \quad \textit{Probabilty of Union}$$

Probability of a Union of Two Events



Not Mutually Exclusive

Proposition

(Probability of a Union of Two Events)

Let $E, F \subseteq \Omega$ be two events from sample space Ω of an experiment. Then:

$$(P5) \quad \mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F) \quad \textit{Probabilty of Union}$$

PROOF: The Venn Diagram indicates that $\mathbb{P}(E) + \mathbb{P}(F)$ counts $\mathbb{P}(E \cap F)$ twice! See the textbook for lower-level proof. □

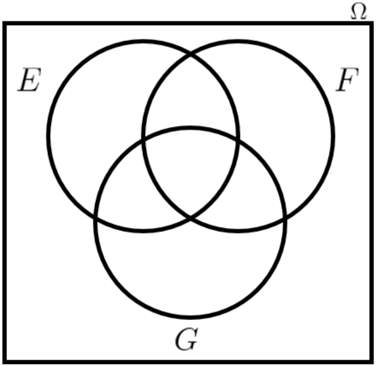
Principle of Inclusion-Exclusion

Proposition

(Principle of Inclusion-Exclusion)

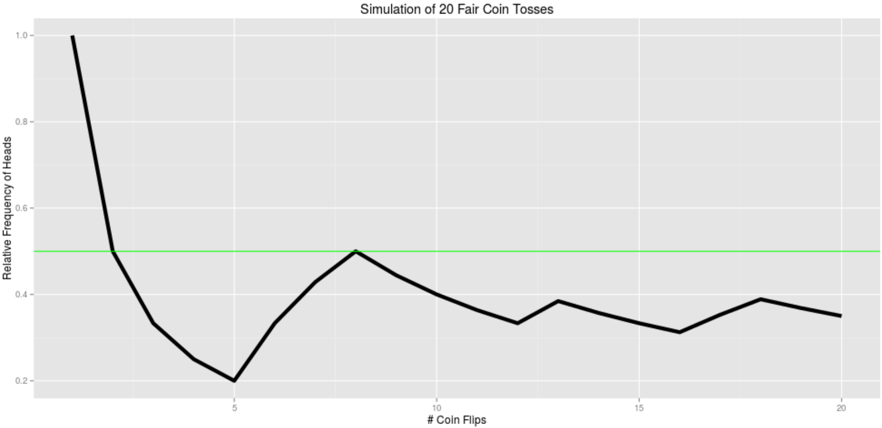
Let $E, F, G \subseteq \Omega$ be three events from sample space Ω of an experiment. Then:

$$\mathbb{P}(E \cup F \cup G) = \mathbb{P}(E) + \mathbb{P}(F) + \mathbb{P}(G) - \mathbb{P}(E \cap F) - \mathbb{P}(E \cap G) - \mathbb{P}(F \cap G) + \mathbb{P}(E \cap F \cap G)$$



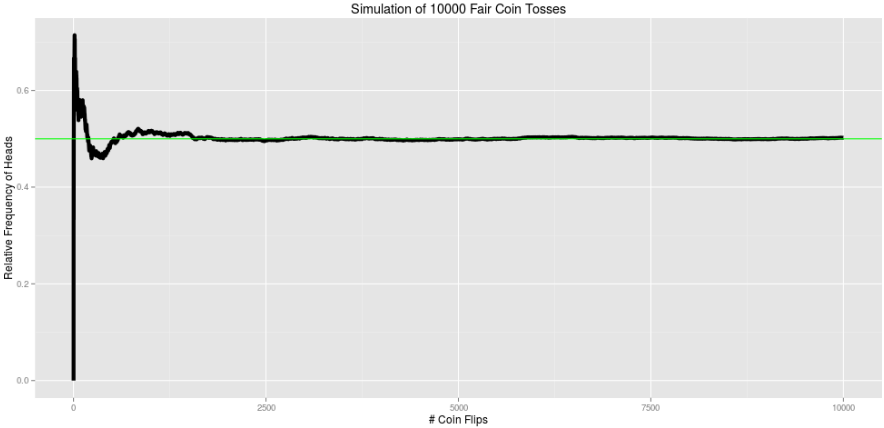
Deep Interpretation of Probability

The axioms & properties do not give a complete interpretation of probability!!
The most intuitive interpretation is to treat probability as a relative frequency:



Deep Interpretation of Probability

The axioms & properties do not give a complete interpretation of probability!!
The most intuitive interpretation is to treat probability as a relative frequency:



Interpretation of Probability

Definition

(Interpretation of Probability)

The **probability of an outcome** in a sample space is the likelihood of the outcome, which is a number between 0 and 1 inclusive.

The **probability of an event** E, denoted $\mathbb{P}(E)$, is the sum of the probabilities of the outcomes that comprise E.

Interpretation of Probability:

$$\begin{array}{rcccccl} & & \mathbb{P}(E) & = & 0 & \Longrightarrow & \text{Event } E \text{ is impossible} \\ 0 & < & \mathbb{P}(E) & < & 0.50 & \Longrightarrow & \text{Event } E \text{ is not likely to occur} \\ & & \mathbb{P}(E) & = & 0.50 & \Longrightarrow & \text{Event } E \text{ has 50-50 chance of occurring} \\ 0.50 & < & \mathbb{P}(E) & < & 1 & \Longrightarrow & \text{Event } E \text{ is likely to occur} \\ & & \mathbb{P}(E) & = & 1 & \Longrightarrow & \text{Event } E \text{ is certain to occur} \end{array}$$

Interpretation of Probability

Definition

(Interpretation of Probability)

The **probability of an outcome** in a sample space is the likelihood of the outcome, which is a number between 0 and 1 inclusive.

The **probability of an event** E, denoted $\mathbb{P}(E)$, is the sum of the probabilities of the outcomes that comprise E.

Examples of Probability:

- "There's a 30% chance of snow tomorrow." $[\mathbb{P}(\text{Snow tomorrow}) = 0.30]$
- "25% of adults get seven hours of sleep." $[\mathbb{P}(\text{7 hrs of sleep}) = 0.25]$
- "All dogs play fetch." $[\mathbb{P}(\text{Playing fetch}) = 1]$
- "There's a 1 in 1000 chance of winning." $\left[\mathbb{P}(\text{Winning}) = \frac{1}{1000}\right]$
- "None of my cats catch mice." $[\mathbb{P}(\text{Catch mice}) = 0]$

Measure of a Set (Definition)

Definition

(Measure of a Set)

The **measure** of a **countable set** is $|E| :=$ (# of elements in E)
The **measure** of a **1D set** is $|E| :=$ (Length of curve E)
The **measure** of a **2D set** is $|E| :=$ (Area of region E)
The **measure** of a **3D set** is $|E| :=$ (Volume of solid E)

The **measure** of the **empty set** is defined to be zero: $|\emptyset| := 0$

Measure of a Countable Set (Example)

Definition

(Measure of a Set)

The **measure** of a **countable set** is $|E| :=$ (# of elements in E)
The **measure** of a **1D set** is $|E| :=$ (Length of curve E)
The **measure** of a **2D set** is $|E| :=$ (Area of region E)
The **measure** of a **3D set** is $|E| :=$ (Volume of solid E)

The **measure** of the **empty set** is defined to be zero: $|\emptyset| := 0$

Example: Let $S = \{\text{Heads}, \text{Tails}\}$.

Then, $|S| =$ (# of elements of S) $= \boxed{2}$

Measure of a 1D Set (Example)

Definition

(Measure of a Set)

The **measure** of a **countable set** is $|E| :=$ (# of elements in E)
The **measure** of a **1D set** is $|E| :=$ (Length of curve E)
The **measure** of a **2D set** is $|E| :=$ (Area of region E)
The **measure** of a **3D set** is $|E| :=$ (Volume of solid E)

The **measure** of the **empty set** is defined to be zero: $|\emptyset| := 0$

Example: Let ℓ be a line segment with length 13.

Then, $|\ell| =$ (Length of ℓ) $= \boxed{13}$

Measure of a 2D Set (Example)

Definition

(Measure of a Set)

The **measure** of a **countable set** is $|E| := (\text{\# of elements in } E)$
The **measure** of a **1D set** is $|E| := (\text{Length of curve } E)$
The **measure** of a **2D set** is $|E| := (\text{Area of region } E)$
The **measure** of a **3D set** is $|E| := (\text{Volume of solid } E)$

The **measure** of the **empty set** is defined to be zero: $|\emptyset| := 0$

Example: Let R be a rectangle with length 2 and width 3.

Then, $|R| = (\text{Area of } R) = (\text{Length}) \times (\text{Width}) = 2 \times 3 = \boxed{6}$

Measure of a 3D Set (Example)

Definition

(Measure of a Set)

The **measure** of a **countable set** is $|E| :=$ (# of elements in E)
The **measure** of a **1D set** is $|E| :=$ (Length of curve E)
The **measure** of a **2D set** is $|E| :=$ (Area of region E)
The **measure** of a **3D set** is $|E| :=$ (Volume of solid E)

The **measure** of the **empty set** is defined to be zero: $|\emptyset| := 0$

Example: Let C be a cube of length 4.

Then, $|C| = (\text{Volume of } C) = (\text{Length})^3 = 4^3 = \boxed{64}$

BE CAREFUL WITH THE VERTICAL BARS!!!

The vertical bars $|\cdot|$ are heavily overloaded in mathematics:

MATHEMATICAL OBJECT	TYPICAL EXPRESSION	MEANING OF VERTICAL BARS
Scalar	$\|a\| = \|-3\|$	Absolute Value of a
Vector	$\|\|\mathbf{v}\|\| = \left\|\left\| \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\|\right\|$	Norm of $\mathbf{v}$
Matrix	$\|A\| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$	Determinant of A
Set	$\|E\| = \|\{1, 2, 3, 4\}\|$	Measure of E

For this course, vectors and matrices will not be considered, so just be mindful of the difference between the absolute value of a scalar & measure of a set.

Probability: Equally Likely Outcomes

Very often, all the outcomes of an experiment are **equally likely to occur**:

Definition

(Probability of an Event)

Let Ω be the sample space of an experiment with **equally likely outcomes**. Let E be an event of the experiment.

Then the **probability** of event E occurring is defined as: $\mathbb{P}(E) = \frac{|E|}{|\Omega|}$

i.e. The probability is the proportion of outcomes that comprise the event.

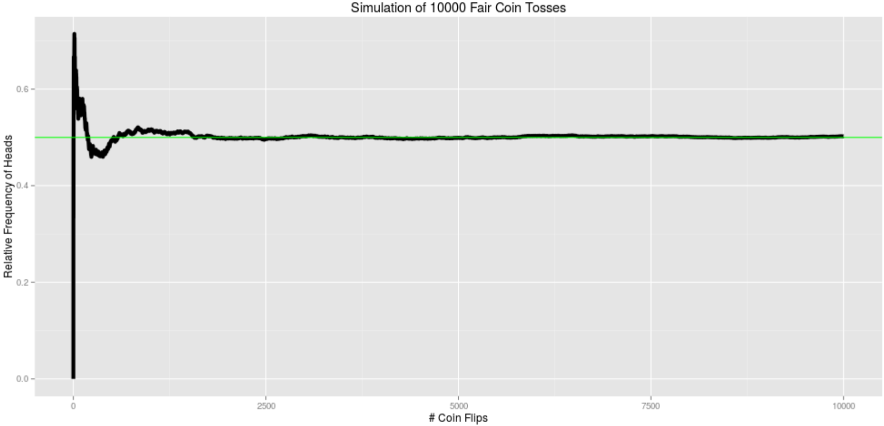
REMARKS:

Often it's impractical to list every outcome of a sample space Ω or event E. When computing probability, only the **measures** of E & Ω are needed.

Fair coins & fair dice always result in equally likely outcomes.

Probability: Equally Likely Outcomes (Example)

Consider the experiment of flipping a coin once with (H)eads and (T)ails.
Let event $E \equiv$ "Heads". Then, sample space $\Omega = \{H, T\}$.



Then: $\mathbb{P}(\text{Heads}) \equiv \mathbb{P}(E) = \dfrac{|E|}{|\Omega|} = \dfrac{|\{H\}|}{|\{H, T\}|} = \dfrac{1}{2} = 0.5$

Textbook Logistics for Section 2.2

- Difference(s) in Terminology:

TEXTBOOK TERMINOLOGY	SLIDES/OUTLINE TERMINOLOGY
Null Event $\emptyset$	Empty Set $\emptyset$
Number of Outcomes in E	Measure of E

- Difference(s) in Notation:

CONCEPT	TEXTBOOK NOTATION	SLIDES/OUTLINE NOTATION
Sample Space	$\mathcal{S}$	Ω
Complement of Event	A'	A^c
Probability of Event	$P(A)$	$\mathbb{P}(A)$
Measure of Event	$N(A)$	$\lvert A\rvert$

Fin

Fin.