EEL 4930 Lecture 7

EEL 4930 Lecture 7

Probability as a Measure of Frequency of Occurrence

EEL 4930 Lecture 7

Probability as a Measure of Frequency of Occurrence

- In the experiments we have conducted, the relative frequencies converge to some constant values when the number of trials gets large

EEL 4930 Lecture 7

Probability as a Measure of Frequency of Occurrence

- In the experiments we have conducted, the relative frequencies converge to some constant values when the number of trials gets large
- The experiments we have conducted are fair or are combined experiments with fair subexperiments, for which we (now) know how to calculate the probabilities

EEL 4930 Lecture 7

Probability as a Measure of Frequency of Occurrence

- In the experiments we have conducted, the relative frequencies converge to some constant values when the number of trials gets large
- The experiments we have conducted are fair or are combined experiments with fair subexperiments, for which we (now) know how to calculate the probabilities
  - the relative frequencies converge to the probabilities

- If the relative frequencies converge for an experiment, we say the experiment possesses *statistical regularity*

- If the relative frequencies converge for an experiment, we say the experiment possesses *statistical regularity* and consider those here:

- If the relative frequencies converge for an experiment, we say the experiment possesses *statistical regularity* and consider those here:

- Consider a random experiment that has K possible outcomes, $K < \infty$

- If the relative frequencies converge for an experiment, we say the experiment possesses *statistical regularity* and consider those here:
- Consider a random experiment that has K possible outcomes, $K < \infty$
- Let $N_k(n) =$ the number of times the outcome is k and let the relative frequency of outcome k be

$$r_k(n) = \frac{N_k(n)}{n}.$$

- If the relative frequencies converge, i.e. $\lim_{n\to\infty} r_k(n)$ is a constant for each k, then:

> For experiments with statistical regularity,
>
> $$\lim_{n\to\infty} r_k(n) = p_k$$
>
> is called the *probability of outcome* k.

Properties of Relative Frequency

Properties of Relative Frequency

- Note that

$$0 \leq N_k(n) \leq n, \quad \forall k$$

Properties of Relative Frequency

- Note that

$$0 \leq N_k(n) \leq n, \quad \forall k$$

because $N_k(n)$ is just the # of times outcome k occurs in n trials.

PROPERTIES OF RELATIVE FREQUENCY

- Note that

$$0 \leq N_k(n) \leq n, \quad \forall k$$

because $N_k(n)$ is just the # of times outcome k occurs in n trials.

Dividing by n yields

$$0 \leq \frac{N_k(n)}{n} = r_k(n) \leq 1, \quad \forall k = 1, \ldots, K \tag{1}$$

- If $1, 2, \ldots, K$ are all of the possible outcomes

- If $1, 2, \ldots, K$ are all of the possible outcomes, then

$$\sum_{k=1}^{K} N_k(n) = n.$$

- If $1, 2, \ldots, K$ are all of the possible outcomes, then

$$\sum_{k=1}^{K} N_k(n) = n.$$

Again, dividing by n yields

$$\sum_{k=1}^{K} r_k(n) = 1. \tag{2}$$

Consider the event E that an even number occurs.

Consider the event E that an even number occurs.

What can we say about the number of times E is observed in n trials?

Consider the event E that an even number occurs.

What can we say about the number of times E is observed in n trials?

$$N_E(n) = N_2(n) + N_4(n) + N_6(n)$$

Consider the event E that an even number occurs.

What can we say about the number of times E is observed in n trials?

$$N_E(n) = N_2(n) + N_4(n) + N_6(n)$$

What have we assumed in developing this equation?

2, 4, and 6 are outcomes & cannot occur at the same time (mutually exclusive)

Then, dividing by n,

$$f_E(n) = \frac{N_E(n)}{n}$$

Then, dividing by n,

$$f_E(n) = \frac{N_E(n)}{n} = \frac{N_2(n) + N_4(n) + N_6(n)}{n}$$

Then, dividing by n,

$$\begin{aligned} f_E(n) &= \frac{N_E(n)}{n} = \frac{N_2(n) + N_4(n) + N_6(n)}{n} \\ &= f_2(n) + f_4(n) + f_6(n) \end{aligned}$$

Then, dividing by n,

$$
\begin{aligned}
f_E(n) &= \frac{N_E(n)}{n} = \frac{N_2(n) + N_4(n) + N_6(n)}{n} \\
&= f_2(n) + f_4(n) + f_6(n)
\end{aligned}
$$

- **General property:**

Then, dividing by n,

$$\begin{aligned} f_E(n) &= \frac{N_E(n)}{n} = \frac{N_2(n)+N_4(n)+N_6(n)}{n} \\ &= f_2(n)+f_4(n)+f_6(n) \end{aligned}$$

- **General property:** If A and B are 2 events that cannot occur simultaneously

Then, dividing by n,

$$\begin{aligned} f_E(n) &= \frac{N_E(n)}{n} = \frac{N_2(n)+N_4(n)+N_6(n)}{n} \\ &= f_2(n)+f_4(n)+f_6(n) \end{aligned}$$

- **General property:** If A and B are 2 events that cannot occur simultaneously, and C is the event that either A or B occurs

Then, dividing by n,

$$
\begin{aligned}
f_E(n) &= \frac{N_E(n)}{n} = \frac{N_2(n)+N_4(n)+N_6(n)}{n} \\
&= f_2(n)+f_4(n)+f_6(n)
\end{aligned}
$$

- **General property:** If A and B are 2 events that cannot occur simultaneously, and C is the event that either A or B occurs, then

$$
f_C(n) = f_A(n) + f_B(n). \tag{3}
$$

- *What are some problems with defining probabilities as the limits of relative frequencies?*

- *What are some problems with defining probabilities as the limits of relative frequencies?*

1. It is not clear when and in what sense the limit exists.

- *What are some problems with defining probabilities as the limits of relative frequencies?*

1. It is not clear when and in what sense the limit exists.
2. It is not possible to perform an experiment an infinite number of times, so the probabilities can never be known exactly.

- *What are some problems with defining probabilities as the limits of relative frequencies?*

1. It is not clear when and in what sense the limit exists.
2. It is not possible to perform an experiment an infinite number of times, so the probabilities can never be known exactly.
3. We cannot use this definition if the experiment cannot be repeated.

We need a *mathematical model of probability* that is not based on a particular application or interpretation.

We need a *mathematical model of probability* that is not based on a particular application or interpretation.

However, any such model should

We need a *mathematical model of probability* that is not based on a particular application or interpretation.

However, any such model should

1. be useful for solving real problems

We need a *mathematical model of probability* that is not based on a particular application or interpretation.
However, any such model should

1. be useful for solving real problems

2. agree with our interpretation of probability as relative frequency

We need a *mathematical model of probability* that is not based on a particular application or interpretation.
However, any such model should

1. be useful for solving real problems
2. agree with our interpretation of probability as relative frequency
3. agree with our intuition

We need a *mathematical model of probability* that is not based on a particular application or interpretation.
However, any such model should

1. be useful for solving real problems
2. agree with our interpretation of probability as relative frequency
3. agree with our intuition (*where appropriate!*)

Probability Spaces

We define a probability space as a mathematical construction containing three elements.

PROBABILITY SPACES

We define a probability space as a mathematical construction containing three elements.We say that a probability space is a *triple*:

$S(\cancel{\Omega}, \mathscr{F}, P)$**.**

PROBABILITY SPACES

We define a probability space as a mathematical construction containing three elements.We say that a probability space is a *triple***:** $(\Omega, \mathscr{F}, P)$**.**

- We have already defined the **sample space** $\cancel{\Omega}$ S and **event class** $\mathscr{F}$

PROBABILITY SPACES

We define a probability space as a mathematical construction containing three elements.We say that a probability space is a ***triple*****:** $(\Omega, \mathscr{F}, P)$**.**

- We have already defined the **sample space** Ω and **event class** $\mathscr{F}$
- We need to specify P:

Probability Spaces

We define a probability space as a mathematical construction containing three elements.We say that a probability space is a *triple*: $(\Omega, \mathscr{F}, P)$**.**

- We have already defined the **sample space** Ω and **event class** $\mathscr{F}$
- We need to specify P:

The *probability measure*, denoted by P

Probability Spaces

We define a probability space as a mathematical construction containing three elements.We say that a probability space is a *triple*: $(\Omega, \mathscr{F}, P)$**.**

- We have already defined the **sample space** Ω and **event class** $\mathscr{F}$
- We need to specify P:

The *probability measure*, denoted by P is a numerically-valued set function that maps all members of $\mathscr{F}$

PROBABILITY SPACES

We define a probability space as a mathematical construction containing three elements.We say that a probability space is a *triple*: $(\Omega, \mathscr{F}, P)$.

- We have already defined the **sample space** Ω and **event class** $\mathscr{F}$

- We need to specify P:

> The *probability measure*, denoted by P is a numerically-valued set function that maps all members of $\mathscr{F}$ onto $\mathbb{R}$.

Axioms of Probability

Axioms: We specify a minimal set of rules that P must obey:

Axioms of Probability

Axioms: We specify a minimal set of rules that P must obey:

I. $\forall E \in \mathscr{F}, P(E) \geq 0$

Axioms of Probability

Axioms: We specify a minimal set of rules that P must obey:

I. $\forall E \in \mathscr{F}, P(E) \geq 0$

II. $P(\overset{S}{\cancel{\Omega}}) = 1$

AXIOMS OF PROBABILITY

Axioms: We specify a minimal set of rules that P must obey:

I. $\forall E \in \mathscr{F}, P(E) \geq 0$

II. $P(\Omega) = 1$

III. $\forall E, F \in \mathscr{F}, P(E \cup F) = P(E) + P(F)$ if $\underbrace{E \cap F = \emptyset}_{\text{m. e.}}$

III. $\forall E, F \in \mathscr{F}, P(E \cup F) = P(E) + P(F)$ if $E \cap F = \emptyset$

III'. If $A_1, A_2, \ldots$ is a sequence of event such that $A_i \cap A_j = \emptyset \; \forall i \neq j$, then

$$P\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} P\left[A_k\right].$$

- We can use Axioms I-III to deal with finite sample spaces. However, Axiom III' is required instead of Axiom III for infinite sample spaces

- Axiom III is a special case of Axiom III'

- The unions and summations in Axiom III' are over *countable* index sets only

Corollaries

Corollaries

Let $A \in \mathscr{F}$ and $B \in \mathscr{F}$. Then the following properties of P can be derived from the axioms and the mathematical structure of $\mathscr{F}$:

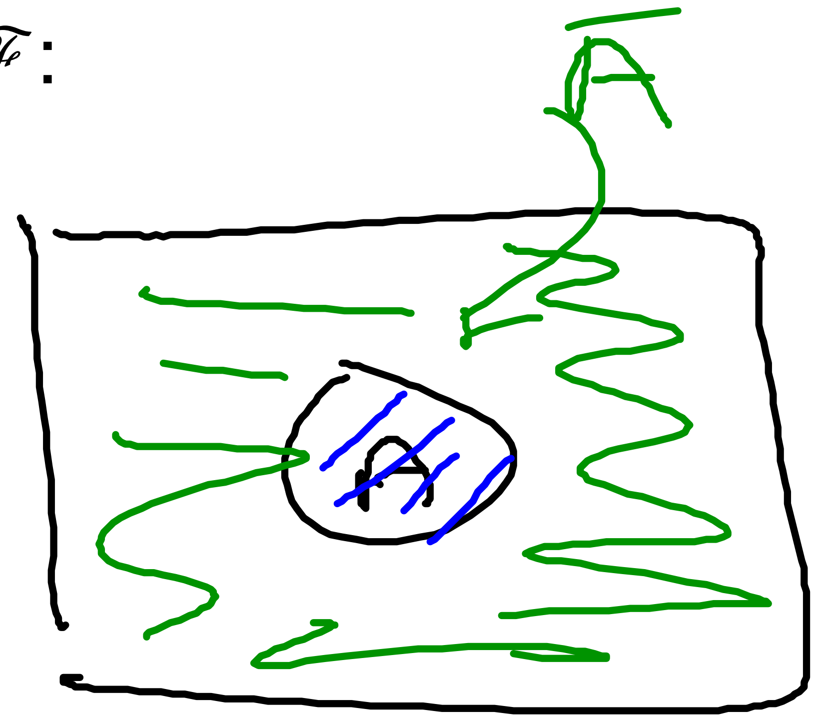
1. $P(A^c) = 1 - P(A)$

$\bar{A}$

$P(S) = 1$

$P(A \cup A^c) = 1$

$P(A) + P(A^c) = 1$



$A \cap A^c = \emptyset$

$A \cup A^c = S$

Corollaries

Let $A \in \mathscr{F}$ and $B \in \mathscr{F}$. Then the following properties of P can be derived from the axioms and the mathematical structure of $\mathscr{F}$:

1. $P(A^c) = 1 - P(A)$

2. $P(A) \leq 1$

$$P(A) = 1 - \underbrace{P(\bar{A})}_{\geq 0}$$

$$\Rightarrow P(A) \leq 1$$

2. $P(A) \leq 1$

3. $P(\emptyset) = 0$

$P(\emptyset) = 1 - P(S) = 1 - 1 = 0$

$\emptyset = \overline{S}$

2. $P(A) \leq 1$

3. $P(\emptyset) = 0$

4. If $A_1, A_2, \ldots, A_n$ are pairwise mutually exclusive, then

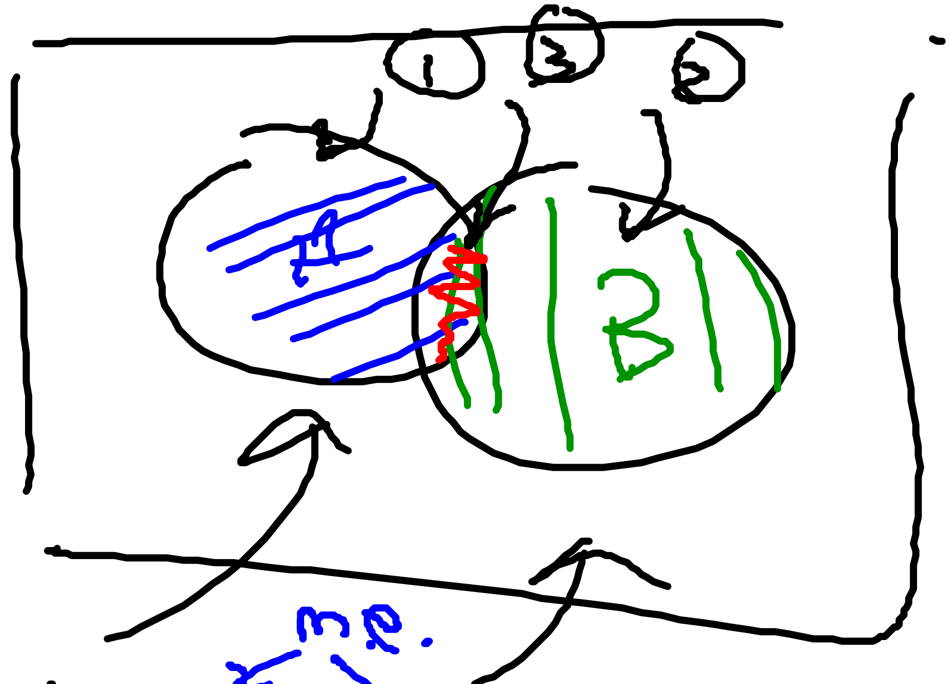
$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k)$$

Proof is by induction

5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= P(A \cap \bar{B}) + P(B \cap \bar{A}) + P(A \cap B)$

$= [P(A) - P(A \cap B)] + [P(B) - P(A \cap B)] + P(A \cap B)$



m.e.

$A = (A \cap B) \cup (A \cap \bar{B})$

m.e.

$B = (B \cap A) \cup (B \cap \bar{A})$

$P(A) = P(A \cap B) + P(A \cap \bar{B})$

$P(B) = P(A \cap B) + P(B \cap \bar{A})$

5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proof:

5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proof:

EX **Take 2:** A fair six-sided die is rolled twice. What is the probability of observing either a 1 or a 2 on the top face on either roll?

Let E_i = 1 or 2 on roll i

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

$$= \frac{1}{3} + \frac{1}{3} - \frac{4}{36}$$

(1,1) (1,2)
(2,1) (2,2)

$$= \frac{2}{6} - \frac{1}{9} = 5/9$$

6.

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_j) - \sum_{j<k} P(A_j \cap A_k) + \cdots$$
$$+(-1)^{(n+1)} P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

Add all single events, subtract off all intersections of pairs of events, add in all intersections of 3 events, ...

6.

$$
\begin{aligned}
P\left(\bigcup_{k=1}^{n} A_k\right) &= \sum_{k=1}^{n} P(A_j) - \sum_{j<k} P(A_j \cap A_k) + \cdots \\
&\quad + (-1)^{(n+1)} P(A_1 \cap A_2 \cap \cdots \cap A_n)
\end{aligned}
$$

Add all single events, subtract off all intersections of pairs of events, add in all intersections of 3 events, ...

Proof is by induction.

7. If $A \subset B$, then $P(A) \leq P(B)$.

7. If $A \subset B$, then $P(A) \leq P(B)$.
Proof: