

# Probability and Statistics

## Understanding the Basics

DRME

December 26, 2023

# The Set of All Possible Events

- ▶ Overview of the concept:
  - ▶ A fundamental concept in probability theory.
  - ▶ Represents all potential outcomes or occurrences in an experiment or situation.
- ▶ Significance in probability theory:
  - ▶ Forms the sample space for a given experiment.
  - ▶ Essential for defining probabilities of specific events.

## Example: Coin Toss

- ▶ Consider the experiment of tossing a fair coin.
- ▶ The set of all possible events:  $\{\text{Heads}, \text{Tails}\}$ .
- ▶ Each outcome in the set is a possible event.

## Example: Rolling a Die

- ▶ Experiment: Rolling a six-sided die.
- ▶ The set of all possible events:  $\{1, 2, 3, 4, 5, 6\}$ .
- ▶ Each face of the die represents a possible event.

# Coin Flips: Why You Should Care

- ▶ Illustration using coin flips.
- ▶ Linking the concept to real-world scenarios.

# Probability: Outside or Inside the Head

- ▶ Outside the head: Long-run relative frequency.
- ▶ Inside the head: Subjective belief.
- ▶ Probabilities assign numbers to possibilities.

# Outside the Head: Long-run Relative Frequency

- ▶ Simulating a long-run relative frequency.
- ▶ Deriving a long-run relative frequency.

# Inside the Head: Subjective Belief

- ▶ Calibrating a subjective belief by preferences.
- ▶ Describing a subjective belief mathematically.



# Probabilities Assign Numbers to Possibilities

- ▶ Probabilities are assigned to different outcomes to quantify uncertainty.
- ▶ The assignment of probabilities is a fundamental concept in probability theory.

# Understanding How Probabilities Are Assigned

- ▶ Probabilities are numerical measures representing the likelihood of events.
- ▶ Assigning probabilities involves assessing the chance of different outcomes.
- ▶ Probabilities are expressed as values between 0 and 1, where 0 indicates impossibility, 1 indicates certainty, and values in between represent varying degrees of likelihood.

## Example: Coin Toss

- ▶ Consider a fair coin toss.
- ▶ There are two possible outcomes: heads (H) or tails (T).
- ▶ Since the coin is fair, the probability of getting heads is 0.5, and the probability of getting tails is also 0.5.

## Example: Rolling a Six-sided Die

- ▶ Suppose you roll a standard six-sided die.
- ▶ Each face has an equal chance of landing face up.
- ▶ The probability of rolling a specific number, say 3, is  $\frac{1}{6}$  because there are six possible outcomes.

## Example: Drawing a Card from a Deck

- ▶ Consider drawing a single card from a standard deck of 52 playing cards.
- ▶ The probability of drawing an Ace is  $\frac{4}{52}$  since there are four Aces in the deck.
- ▶ The probability of drawing a red card is  $\frac{26}{52}$  since half of the cards are red.

# Probability Distributions

- ▶ Overview of Probability Distributions [\[\[see distributions.pdf\]\]](#).

# Discrete Distributions: Probability Mass

- ▶ Definition: Probability mass function for discrete distributions.
- ▶ Examples:
  - ▶ Bernoulli distribution.
  - ▶ Binomial distribution.
  - ▶ Poisson distribution.

# Continuous Distributions: Rendezvous with Density

- ▶ Definition: Probability density function for continuous distributions.
- ▶ Examples:
  - ▶ Uniform distribution.
  - ▶ Exponential distribution.
  - ▶ Normal distribution.



# Appendix A

## Probability Distributions and Conjugate Priors

TABLE A.1

Probability Distributions Used in Ecological Modeling to Represent Stochasticity in Discrete Random Variables ( $z$ )

Distribution	Random variable	Parameters	Moments
Poisson $[z \lambda] = \frac{\lambda^z e^{-\lambda}}{z!}$	Counts of things that occur randomly over time or space, e.g., the number of birds in a forest stand, the number of fish in a kilometer of river, the number of prey captured per minute	$\lambda$ , the mean number of occurrences per time or space $\lambda = \mu$	$\mu = \lambda$ $\sigma^2 = \lambda$
Binomial $[z \eta, \phi] = \binom{\eta}{z} \phi^z (1-\phi)^{\eta-z}$ $\binom{\eta}{z} = \frac{\eta!}{z!(\eta-z)!}$	Number of "successes" on a given number of trials, e.g., number of survivors in a sample of individuals, number of plots containing an exotic species from a sample, number of terrestrial pixels that are vegetated in an image	$\eta$ , the number of trials $\phi$ , the probability of a success $\phi = 1 - \sigma^2/\mu$ $\eta = \mu^2/(\mu - \sigma^2)$	$\mu = \eta\phi$ $\sigma^2 = \eta\phi(1-\phi)$
Bernoulli $[z \phi] = \phi^z (1-\phi)^{1-z}$	A special case of the binomial where the number of trials= 1 and the random variable can take on values 0 or 1; widely used in survival analysis, occupancy models	$\phi$ , the probability that the random variable= 1 $\phi = \mu$ $\phi = 1/2 + 1/2\sqrt{1-4\sigma^2}$	$\mu = \phi$ $\sigma^2 = \phi(1-\phi)$
Negative binomial $[z \lambda, \kappa] = \frac{\Gamma(\kappa+\lambda)}{\Gamma(\kappa)\Gamma(\lambda)} \left(\frac{\kappa}{\kappa+\lambda}\right)^\kappa \times \left(\frac{\lambda}{\kappa+\lambda}\right)^z$	Counts of things occurring randomly over time or space, as with the Poisson; includes dispersion parameter $\kappa$ allowing the variance to exceed the mean	$\lambda$ , the mean number of occurrences per time or space $\kappa$ , the dispersion parameter $\lambda = \mu$ $\kappa = \mu^2/(\sigma^2 - \mu)$	$\mu = \lambda$ $\sigma^2 = \lambda + \lambda^2/\kappa$

**TABLE A.1**  
(continued)

Distribution	Random variable	Parameters	Moments
Multinomial $[z   \eta, \phi] = \eta! \prod_{i=1}^k \frac{\phi_i^{z_i}}{z_i!}$	Counts that fall into $k > 2$ categories, e.g., number of individuals in age classes, number of pixels in different landscape categories, number of species in trophic categories in a sample from a food web	$\mathbf{z}$ , a vector giving the number of counts in each category $\phi$ , a vector of the probabilities of occurrence in each category $\sum_{i=1}^k \phi_i = 1$ $\sum_{i=1}^k z_i = \eta$	$\mu_i = \eta \phi_i$ $\sigma_i^2 = \eta \phi_i (1 - \phi_i)$

Note: We use  $\mu$  to symbolize the first moment of the distribution,  $\mu = E(z)$ , and  $\sigma^2$  to symbolize the second central moment,  $\sigma^2 = E((z - \mu)^2)$ .

**TABLE A.2**  
Probability distributions Used in Ecological Modeling to Represent Stochasticity in Continuous Random Variables ( $z$ )

Continuous Distributions	Random variable ( $z$ )	Parameters	Moments
Normal $[z   \mu, \sigma^2] = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$	Continuously distributed quantities that can take on positive or negative values; sums of things.	$\mu, \sigma^2$	$\mu, \sigma^2$
Lognormal $[z   \alpha, \beta] = \frac{1}{z \sqrt{2\pi\beta^2}} e^{-\frac{(\log(z) - \alpha)^2}{2\beta^2}}$	Continuously distributed quantities with nonnegative values. Random variables with the property that their logs are normally distributed. Thus, if $z$ is normally distributed, then $\exp(z)$ is lognormally distributed. Represents products of things. The variance increases with the mean squared.	$\alpha$ , the mean of $z$ on the log scale $\beta$ , the standard deviation of $z$ on the log scale $\alpha = \log(\text{median}(z))$ $\alpha = \log(\mu) - \frac{1}{2} \log\left(\frac{\sigma^2 + \mu^2}{\mu^2}\right)$ $\beta = \sqrt{\log\left(\frac{\sigma^2 + \mu^2}{\mu^2}\right)}$	$\mu = e^{\alpha + \frac{\beta^2}{2}}$ $\text{median}(z) = e^{\alpha}$ $\sigma^2 = (e^{\beta^2} - 1)e^{2\alpha + \beta^2}$

**TABLE A.2**  
(continued)

Continuous Distributions	Random variable ( $z$ )	Parameters	Moments
Gamma $[z \alpha, \beta] = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\beta z}$ $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$	The time required for a specified number of events to occur in a Poisson process; any continuous quantity that is nonnegative.	$\alpha$ = shape $\beta$ = rate $\alpha = \frac{\mu}{\sigma^2}$ $\beta = \frac{\mu}{\sigma^2}$ Note—be very careful about rate, defined as above, and scale $= \frac{1}{\beta}$ .	$\mu = \frac{\alpha}{\beta}$ $\sigma^2 = \frac{\alpha}{\beta^2}$
Inverse gamma $[z \alpha, \beta] = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{-\alpha-1} \exp\left(-\frac{\beta}{z}\right)$	The reciprocal of a gamma-distributed random variable.	$\alpha$ = shape $\beta$ = scale $\alpha = \frac{\mu^2}{\sigma^2} + 2$ $\beta = \mu \left( \frac{\mu^2}{\sigma^2} + 1 \right)$	$\mu = \frac{\beta}{\alpha-1}$ for $\alpha > 1$ $\sigma^2 = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$
Exponential $[z \alpha, \beta] = \lambda e^{-\lambda z}$	Intervals of time between sequential events that occur randomly over time or space. If the number of events is Poisson distributed, then the times between events are exponentially distributed.	$\lambda$ , the mean number of occurrences per time or space $\lambda = \frac{1}{\mu}$	$\mu = \frac{1}{\lambda}$ $\sigma^2 = \left(\frac{1}{\lambda}\right)^2$
Beta $[z \alpha, \beta] = \frac{B}{\Gamma(\alpha)\Gamma(\beta)} z^{\alpha-1} (1-z)^{\beta-1}$ $B = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ Because $B$ is a normalizing constant, $[z \alpha, \beta] \propto z^{\alpha-1} (1-z)^{\beta-1}$	Continuous random variables that can take on values between 0 and 1, any random variable that can be expressed as a proportion; survival; proportion of landscape invaded by exotic; probabilities of transition from one state to another.	$\alpha = \frac{(\mu^2 - \mu^3 - \mu\sigma^2)}{\sigma^2}$ $\beta = \frac{(\mu - 2\mu^2 + \mu^3 - \sigma^2 + \mu\sigma^2)}{\sigma^2}$	$\mu = \frac{\alpha}{\alpha + \beta}$ $\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$
Dirichlet $[z \alpha] = \frac{\prod_{j=1}^k z_j^{\alpha_j-1}}{\Gamma(\alpha_j)}$	Vectors of more than two elements of continuous random variables that can take on values between 0 and 1 and that sum to 1.	$\alpha_i = \mu_i \alpha_0$ $\alpha_0 = \sum_{i=1}^k \alpha_i$	$\mu_i = \frac{\alpha_i}{\sum_{i=1}^k \alpha_i}$ $\sigma_i^2 = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}$

# Properties of Probability Density Functions

- ▶ Understanding the properties that characterize probability density functions.
- ▶ Emphasis on normalization and non-negativity.

# The Normal Probability Density Function

- ▶ Special focus on the normal distribution.
- ▶ Shape, mean, and standard deviation.
- ▶ Real-world examples: Height, IQ scores.

# Mean and Variance of a Distribution

- ▶ Mean as minimized variance.

# Highest Density Interval (HDI)

- ▶ Understanding the concept of HDI.

# Concept of HDI

- ▶ The Highest Density Interval (HDI) is a statistical concept used in probability distributions.
- ▶ It provides a range of values within which a specified portion of the probability density function lies.



# Why HDI Matters

- ▶ HDI is valuable for summarizing uncertainty about a parameter.
- ▶ It's particularly useful when dealing with complex distributions or posterior distributions from Bayesian analysis.

# Calculating HDI

- ▶ HDI is often calculated numerically using methods such as Markov Chain Monte Carlo (MCMC).
- ▶ It represents the narrowest interval that contains a certain predefined probability mass.

# Interpretation of HDI

- ▶ The width of the HDI reflects the precision of our knowledge about the parameter.
- ▶ A narrow HDI indicates more precise estimation, while a wider HDI suggests greater uncertainty.

# Two-Way Distributions

- ▶ Conditional Probability
- ▶ Independence of Attributes

# Conditional Probability

Conditional probability is the probability of an event occurring given that another event has already occurred. It is denoted by  $P(A|B)$ , representing the probability of event A given that event B has occurred. The formula for conditional probability is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

where  $P(A \cap B)$  is the probability of both events A and B occurring, and  $P(B)$  is the probability of event B occurring.

# Independence of Attributes

Two events, A and B, are considered independent if the occurrence or non-occurrence of one event does not affect the probability of the other event. Mathematically, events A and B are independent if:

$$P(A \cap B) = P(A) \cdot P(B)$$

In other words, the joint probability of A and B equals the product of their individual probabilities. If this equation holds, A and B are independent; otherwise, they are dependent.

# Conditional Probability

- ▶ Conditional Probability:  $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- ▶ Example:
  - ▶ Suppose you have a deck of cards. Let  $A$  be the event of drawing a red card, and  $B$  be the event of drawing a heart. The conditional probability of drawing a red card given that it is a heart is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{2}}{\frac{1}{4}} = \frac{2}{1} = 2$$

# Independence of Attributes

- ▶ Independence of Attributes:  $P(A \cap B) = P(A) \cdot P(B)$
- ▶ Example:
  - ▶ Consider two events:  $C$  is the event of rolling a 4 on a six-sided die, and  $D$  is the event of getting heads on a fair coin toss. If  $C$  and  $D$  are independent, then:

$$P(C \cap D) = P(C) \cdot P(D) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$$



# Joint Probability

- ▶ Definition: The probability of the occurrence of two or more events simultaneously.
- ▶ Denoted as  $P(A \cap B)$  for events  $A$  and  $B$ .

# Joint Probability Example

- ▶ Consider rolling a six-sided die.
- ▶ Let  $A$ : The event of rolling an even number.
- ▶ Let  $B$ : The event of rolling a number greater than 3.
- ▶ Find  $P(A \cap B)$ .

# Calculating Joint Probability

- ▶ Using the formula:  $P(A \cap B) = P(A) \times P(B|A)$
- ▶  $P(A)$ : Probability of event  $A$
- ▶  $P(B|A)$ : Probability of event  $B$  given that event  $A$  has occurred.

# Joint Probability Example

- ▶  $P(A)$ : Probability of rolling an even number  $= \frac{3}{6} = \frac{1}{2}$
- ▶  $P(B|A)$ : Probability of rolling a number greater than 3 given that an even number is rolled  $= \frac{2}{3}$
- ▶  $P(A \cap B) = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$

# Joint Probability Interpretation

- ▶ Joint probability provides a measure of the likelihood of multiple events occurring together.
- ▶ Important in understanding relationships between events.

# Applications

- ▶ Finance: Probability of a stock both gaining value and exceeding a certain threshold.
- ▶ Medicine: Probability of a patient having multiple symptoms simultaneously.
- ▶ Weather: Probability of rain and high winds occurring together.

# Rule 1: Product Rule

## Product Rule

$$P(A \cap B) = P(A|B) \cdot P(B)$$

- Interpretation: Probability of both events A and B occurring.

# Example: Genetic Inheritance

- ▶ Event A: Offspring having a specific genetic trait.
- ▶ Event B: Parent carrying the gene for the trait.
- ▶ Using the product rule to calculate the joint probability.



# Example: Species Coexistence in Ecology

- ▶ Event A: Presence of species X in an ecosystem.
- ▶ Event B: Availability of a specific environmental condition.
- ▶ Applying the product rule to understand coexistence probabilities.

## Rule 2: Chain Rule

### Chain Rule Formula

$$P(A \cap B \cap C) = P(A|B \cap C) \cdot P(B|C) \cdot P(C)$$

- ▶ The formula for three events, generalizable to more variables.

# Application in Biology

## ► Example 1: Gene Expression

- $A$ : Gene activation,  $B$ : Cellular environment,  $C$ : External signals.
- Probability of gene activation influenced by the cellular environment **AND** external signals.

## ► Example 2: Ecosystem Dynamics

- $A$ : Predation occurrence,  $B$ : Prey abundance,  $C$ : Environmental conditions.
- Probability of predation depends on prey abundance **AND** environmental conditions.

# Generalization

- ▶ Rule 2 can be generalized to more variables.
- ▶ **Example 3: Evolutionary Processes**
  - ▶  $A$ : Adaptation,  $B$ : Genetic variation,  $C$ : Selection pressure.
  - ▶ Probability of adaptation influenced by genetic variation **AND** selection pressure.

## Rule 3: Marginalization

### Marginal Probability

$$P(A) = \sum_B P(A \cap B)$$

Interpretation and practical implications.

## Example: Marginalization

- ▶ Consider a joint probability distribution table representing the occurrence of two traits in a population.
- ▶ Traits: Trait A (dominant/recessive) and Trait B (present/absent).
- ▶ Joint probabilities are given in the table.

# Joint Probability Distribution Table

	Trait B Present	Trait B Absent
Trait A Dominant	$P(A, B)$	$P(A, \neg B)$
Trait A Recessive	$P(\neg A, B)$	$P(\neg A, \neg B)$

# Calculation of Marginal Probabilities

- ▶ Marginal probability of Trait A:  $P(A) = P(A, B) + P(A, \neg B)$
- ▶ Marginal probability of Trait B:  $P(B) = P(A, B) + P(\neg A, B)$



# Calculation of Marginal Probabilities

- ▶ TABLE 3.1 [[@hobbs2015]]

# Interpretation

- ▶ Marginalization allows us to analyze the probability of individual traits independently.
- ▶ This is crucial for understanding the genetic composition of a population.