Probability and Likelihood

CCBB Introduction to Biostats

September 11, 2015

Recipe for a probability model

1. Define the mutually exclusive outcomes of all random variables (aka the sample space).

Recipe for a probability model

1. Define the mutually exclusive outcomes of all random variables (aka the sample space).
2. Assign probabilities to outcomes (must add to 1).

Recipe for a probability model

1. Define the mutually exclusive outcomes of all random variables (aka the sample space).
2. Assign probabilities to outcomes (must add to 1).
3. Define dependencies among random variables.

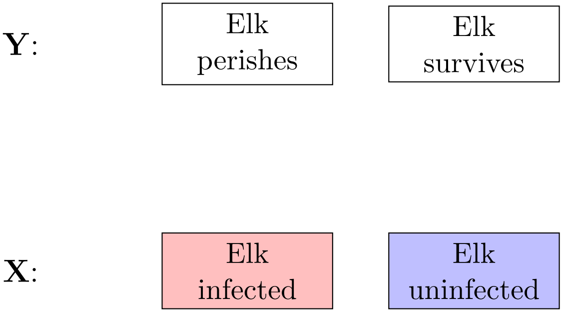
Example probability model: sick elk

Two outcomes each for random variables **X** and **Y**.

Y:	Elk perishes	Elk survives
X:	Elk infected	Elk uninfected

Example probability model: sick elk

Incidence of bovine tuberculosis in the population is 7%.

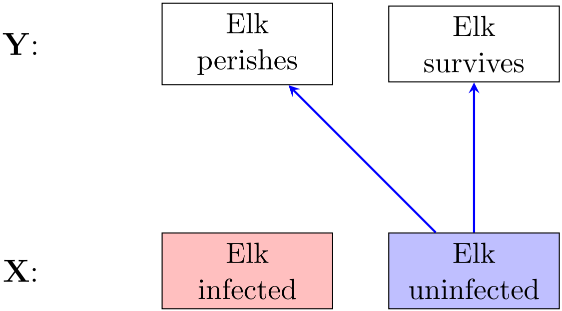


Marginal probability of variable **X** (infection status):

$$\Pr(\text{X} = \text{Infected}) = 0.07$$
$$\Pr(\text{X} = \text{Uninfected}) = 0.93$$

Example probability model: sick elk

Uninfected individuals are more likely to live than to die.

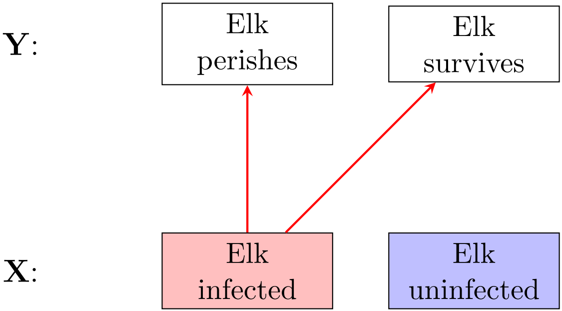


Conditional prob. of variable **Y** (survival) given **X = uninfected**:

$$\Pr(Y = \text{Survives}|X = \text{Uninfected}) = 0.77$$
$$\Pr(Y = \text{Perishes}|X = \text{Uninfected}) = 0.23$$

Example probability model: sick elk

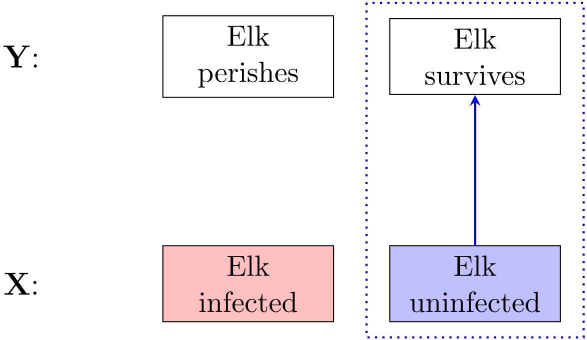
Infected individuals are more likely to die than to live.



Conditional prob. of variable **Y** (survival) given **X = infected**:

$$\Pr(Y = \text{Survives} | X = \text{Infected}) = 0.17$$
$$\Pr(Y = \text{Perishes} | X = \text{Infected}) = 0.83$$

Example probability model: sick elk



Joint prob. of **Y = survival** AND **X = uninfected**:

$$\Pr(\text{Y} = \text{Survives},\ \text{X} = \text{Uninfected}) =$$
$$\Pr(\text{X} = \text{Uninfected}) \times \Pr(\text{Y} = \text{Survives} | \text{X} = \text{Uninfected}) =$$
$$0.93 \times 0.77 = 0.716$$

Example probability model: sick elk

What about the marginal probability of random variable $\mathbf{Y}$?

Two possible ways for outcome $\mathbf{Y} = \textbf{survive}$.

Example probability model: sick elk

What about the marginal probability of random variable **Y**?

Two possible ways for outcome **Y = survive**.

1. **X = uninfected** and **Y = survive**.

The marginal probability equals:

$$\Pr(\mathrm{Y} = \mathrm{Survive}) =$$

$$\Pr(\mathrm{Y} = \mathrm{Survive},\ \mathrm{X} = \mathrm{Uninfected})$$

Example probability model: sick elk

What about the marginal probability of random variable $\mathbf{Y}$?

Two possible ways for outcome $\mathbf{Y} = \mathbf{survive}$.

1. $\mathbf{X} = \mathbf{uninfected}$ and $\mathbf{Y} = \mathbf{survive}$.
2. $\mathbf{X} = \mathbf{infected}$ and $\mathbf{Y} = \mathbf{survive}$.

The marginal probability equals:

$$\Pr(\mathrm{Y} = \mathrm{Survive}) =$$
$$\Pr(\mathrm{Y} = \mathrm{Survive},\ \mathrm{X} = \mathrm{Uninfected}) + \Pr(\mathrm{Y} = \mathrm{Survive},\ \mathrm{X} = \mathrm{Infected})$$

Example probability model: sick elk

What about the marginal probability of random variable **Y**?

Two possible ways for outcome **Y = survive**.

1. **X = uninfected** and **Y = survive**.
2. **X = infected** and **Y = survive**.

The marginal probability equals:

$$\Pr(\mathrm{Y} = \mathrm{Survive}) =$$
$$\Pr(\mathrm{Y} = \mathrm{Survive},\ \mathrm{X} = \mathrm{Uninfected}) + \Pr(\mathrm{Y} = \mathrm{Survive},\ \mathrm{X} = \mathrm{Infected})$$
$$= (0.93 \times 0.77) + (0.07 \times 0.17) = 0.728$$

The sum of values of the joint distribution where **Y = survive**.

Marginal, joint probability as a table

	Infected	**Uninfected**	
Survives	Pr(Infected)× Pr(Survives|Infected)	Pr(Uninfected)× Pr(Survives|Uninfected)	Pr(Survives)
Perishes	Pr(Infected)× Pr(Perishes|Infected)	Pr(Uninfected)× Pr(Perishes|Uninfected)	Pr(Perishes)
	Pr(Infected)	Pr(Uninfected)	1

Marginal, joint probability as text

Marginal probability is the probability of a random variable taking a value, *regardless* of the values of other random variables.

Marginal, joint probability as text

Marginal probability is the probability of a random variable taking a value, *regardless* of the values of other random variables.

Conditional probability is the probability of a random variable taking a value, *given* the values of other random variables.

Marginal, joint probability as text

Marginal probability is the probability of a random variable taking a value, *regardless* of the values of other random variables.

Conditional probability is the probability of a random variable taking a value, *given* the values of other random variables.

Joint probability is the probability of a *combination* of values of several random variables.

Marginal, joint probability as text

Marginal probability is the probability of a random variable taking a value, *regardless* of the values of other random variables.

Conditional probability is the probability of a random variable taking a value, *given* the values of other random variables.

Joint probability is the probability of a *combination* of values of several random variables.

$$\Pr(X,\ Y) = \Pr(X) \times \Pr(Y|X) = \Pr(Y) \times \Pr(X|Y)$$

Marginal, joint probability as text

Marginal probability is the probability of a random variable taking a value, *regardless* of the values of other random variables.

$$\Pr(\mathrm{Y}) = \sum_{x \in \chi} \Pr(\mathrm{Y},\ \mathrm{X} = \mathrm{x}) = \sum_{x \in \chi} \Pr(\mathrm{X} = \mathrm{x}) \times \Pr(\mathrm{Y}|\mathrm{X} = \mathrm{x})$$

Conditional probability is the probability of a random variable taking a value, *given* the values of other random variables.

Joint probability is the probability of a *combination* of values of several random variables.

$$\Pr(\mathrm{X},\ \mathrm{Y}) = \Pr(\mathrm{X}) \times \Pr(\mathrm{Y}|\mathrm{X}) = \Pr(\mathrm{Y}) \times \Pr(\mathrm{X}|\mathrm{Y})$$

Generalizing to other types of RVs

The set of possible outcomes for a random variable can be:

- Infinite (unbounded) or finite (bounded)
- Integer, real, or categorical valued

Generalizing to other types of RVs

The set of possible outcomes for a random variable can be:

- Infinite (unbounded) or finite (bounded)
- Integer, real, or categorical valued

We use functions to distribute (allocate) the probability among outcomes.

Generalizing to other types of RVs

The set of possible outcomes for a random variable can be:

- Infinite (unbounded) or finite (bounded)
- Integer, real, or categorical valued

We use functions to distribute (allocate) the probability among outcomes.

These functions are called probability distribution functions:

- aka PDFs
- with finite number of parameters
- parameters often have an intuitive interpretation
- these functions **must** sum (integrate) to 1.

Shiny time

1. Examples of probability distribution functions.
2. How does conditional/marginal/joint probability translate to continuous RVs?

Code on GitHub:

`https://github.com/sjfox/CCBB_Intro_Biostats/tree/master/week_2`

How likelihood makes our lives better

Problem:

Have data. Need explanation.

How likelihood makes our lives better

Problem:

Have data. Need explanation.

Solution:

How likelihood makes our lives better

Problem:

Have data. Need explanation.

Solution:

1. (Arbitrarily?) choose distribution of data, such that parameters reflect something of interest.

How likelihood makes our lives better

Problem:

Have data. Need explanation.

Solution:

1. (Arbitrarily?) choose distribution of data, such that parameters reflect something of interest.
2. Calculate joint probability of data, conditional on parameter values. This is the **likelihood**.

How likelihood makes our lives better

Problem:

Have data. Need explanation.

Solution:

1. (Arbitrarily?) choose distribution of data, such that parameters reflect something of interest.
2. Calculate joint probability of data, conditional on parameter values. This is the **likelihood**.
3. Vary parameter values, recalculate likelihood.

How likelihood makes our lives better

Problem:

Have data. Need explanation.

Solution:

1. (Arbitrarily?) choose distribution of data, such that parameters reflect something of interest.
2. Calculate joint probability of data, conditional on parameter values. This is the **likelihood**.
3. Vary parameter values, recalculate likelihood.
4. Find parameters which maximize joint probability of data. This is **the maximum likelihood estimate**.

How likelihood makes our lives better

Problem:

Have data. Need explanation.

Solution:

1. (Arbitrarily?) choose distribution of data, such that parameters reflect something of interest.
2. Calculate joint probability of data, conditional on parameter values. This is the **likelihood**.
3. Vary parameter values, recalculate likelihood.
4. Find parameters which maximize joint probability of data. This is **the maximum likelihood estimate**.

Likelihood measures how probable the data are at a parameter value.

Shiny time

1. Examples of calculating the likelihood.

Code on GitHub:

`https://github.com/sjfox/CCBB_Intro_Biostats/tree/master/week_2`

The **likelihood** of some parameter values is the joint probability of the data, conditional on given parameter values.

The **likelihood** of some parameter values is the joint probability of the data, conditional on given parameter values.

$$\mathcal{L}(\text{parameters}|\text{data}) = \prod_{i=1}^{n} \Pr(\text{datum}_i|\text{parameters})$$

The **likelihood** of some parameter values is the joint probability of the data, conditional on given parameter values.

$$\mathcal{L}(\text{parameters}|\text{data}) = \prod_{i=1}^{n} \Pr(\text{datum}_i|\text{parameters})$$

The **maximum likelihood estimate** (MLE) is the parameter value which maximizes the joint probability of the data.

The **likelihood** of some parameter values is the joint probability of the data, conditional on given parameter values.

$$\mathcal{L}(\text{parameters}|\text{data}) = \prod_{i=1}^{n} \Pr(\text{datum}_i|\text{parameters})$$

The **maximum likelihood estimate** (MLE) is the parameter value which maximizes the joint probability of the data.

The **likelihood** of some parameter values is the joint probability of the data, conditional on given parameter values.

$$\mathcal{L}(\text{parameters}|\text{data}) = \prod_{i=1}^{n} \Pr(\text{datum}_i|\text{parameters})$$

The **maximum likelihood estimate** (MLE) is the parameter value which maximizes the joint probability of the data.

The **Fisher information** of the parameters is the negative of the **Hessian** (the second derivative of the log-likelihood).

The **likelihood** of some parameter values is the joint probability of the data, conditional on given parameter values.

$$\mathcal{L}(\text{parameters}|\text{data}) = \prod_{i=1}^{n} \Pr(\text{datum}_i|\text{parameters})$$

The **maximum likelihood estimate** (MLE) is the parameter value which maximizes the joint probability of the data.

The **Fisher information** of the parameters is the negative of the **Hessian** (the second derivative of the log-likelihood).

$$\mathcal{I}(\text{parameters}) = -\frac{\partial^2 \ln \mathcal{L}(\text{parameters}|\text{data})}{\partial(\text{parameters})^2}$$

The **likelihood** of some parameter values is the joint probability of the data, conditional on given parameter values.

$$\mathcal{L}(\text{parameters}|\text{data}) = \prod_{i=1}^{n} \Pr(\text{datum}_i|\text{parameters})$$

The **maximum likelihood estimate** (MLE) is the parameter value which maximizes the joint probability of the data.

The **Fisher information** of the parameters is the negative of the **Hessian** (the second derivative of the log-likelihood).

$$\mathcal{I}(\text{parameters}) = -\frac{\partial^2 \ln \mathcal{L}(\text{parameters}|\text{data})}{\partial(\text{parameters})^2}$$

Measures the information the data bring about the parameters, and is used to calculate the **standard error** of the MLE.

Takeaway:

- Types of probability and how they relate.
- What a probability distribution function is.
- What likelihood is, and how to calculate it.
- What the likelihood tells us about the possible parameter values that generated our data.
- How uncertainty in the maximum likelihood estimate is measured (with the Fisher information/Hessian)