Probability Practice

Data Science: Jordan Boyd-Graber
University of Maryland
FEBRUARY 12, 2018

Big Picture

- Probabilities
- Need to have intuitions for later models
- Key ideas: marginal distributions, independence

Marginal Probabilities

- A voter can either be a Democrat or Republican (f) and has an age (A)
  - $p(F = \text{D}) = .45$
  - $p(A < 30, F = \text{D}) = .2$, $p(A < 30, F = \text{R}) = .1$
  - $p(A > 50, F = \text{D}) = .1$
  - $p(30 \leq a \leq 50) = .3$

Marginal Probabilities

- A voter can either be a Democrat or Republican (f) and has an age (A)
  - $p(F = \text{D}) = .45$
  - $p(A < 30, F = \text{D}) = .2$, $p(A < 30, F = \text{R}) = .1$
  - $p(A > 50, F = \text{D}) = .1$
  - $p(30 \leq a \leq 50) = .3$
- What is $p(30 \leq A \leq 50, F = \text{D})$?
- What is $p(30 \leq A \leq 50, F = \text{R})$?
- What is $p(A > 50, F = \text{R})$?

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	
$30 \leq a \leq 50$			.3
> 50	.1		
Marginal	.45		1.0

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	.3
$30 \leq a \leq 50$			.3
> 50	.1		
Marginal	.45		1.0

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	.3
$30 \leq a \leq 50$			.3
> 50	.1		.4
Marginal	.45	.55	1.0

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	.3
$30 \leq a \leq 50$	x		.3
> 50	.1		.4
Marginal	.45	.55	1.0

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	.3
$30 \leq a \leq 50$	x	y	.3
> 50	.1		.4
Marginal	.45	.55	1.0

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	.3
$30 \leq a \leq 50$	x	y	.3
> 50	.1	z	.4
Marginal	.45	.55	1.0

$$.2 + x + .1 = .45$$
$$x + y = .3$$
$$.1 + z = .4$$

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	.3
$30 \leq a \leq 50$	.15	y	.3
> 50	.1	z	.4
Marginal	.45	.55	1.0

$$.2 + x + .1 = .45$$
$$x + y = .3$$
$$.1 + z = .4$$

x = .45 - .1 - .2 = .15

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	.3
$30 \leq a \leq 50$	.15	.15	.3
> 50	.1	z	.4
Marginal	.45	.55	1.0

$$.2 + x + .1 = .45$$

$$x + y = .3$$

$$.1 + z = .4$$

y = .3 - x = .3 - .15 = .15

Solving the Marginal Probabilities

	D	R	Marginal
< 30	.2	.1	.3
$30 \leq a \leq 50$	.15	.15	.3
> 50	.1	.3	.4
Marginal	.45	.55	1.0

$$.2 + x + .1 = .45$$
$$x + y = .3$$
$$.1 + z = .4$$

z = .4 - .1 = .3

What if age and party were independent?

	D	R	Marginal
< 30			.3
$30 \leq a \leq 50$			.3
> 50			.4
Marginal	.45	.55	1.0

What if age and party were independent?

	D	R	Marginal
< 30	.135		.3
$30 \leq a \leq 50$			.3
> 50			.4
Marginal	.45	.55	1.0

What if age and party were independent?

	D	R	Marginal
< 30	.135	.165	.3
$30 \leq a \leq 50$			.3
> 50			.4
Marginal	.45	.55	1.0

What if age and party were independent?

	D	R	Marginal
< 30	.135	.165	.3
$30 \leq a \leq 50$	.135		.3
> 50			.4
Marginal	.45	.55	1.0

What if age and party were independent?

	D	R	Marginal
< 30	.135	.165	.3
$30 \leq a \leq 50$	.135	.165	.3
> 50			.4
Marginal	.45	.55	1.0

What if age and party were independent?

	D	R	Marginal
< 30	.135	.165	.3
$30 \leq a \leq 50$	.135	.165	.3
> 50	.18		.4
Marginal	.45	.55	1.0

What if age and party were independent?

	D	R	Marginal
< 30	.135	.165	.3
$30 \leq a \leq 50$	.135	.165	.3
> 50	.18	.22	.4
Marginal	.45	.55	1.0

Expected Value

In Las Vegas the roulette wheel has a 0 and a 00 and then the numbers 1 to 36 marked on equal slots; the wheel is spun and a ball stops randomly in one slot. When a player bets 1 dollar on a number, he receives 36 dollars if the ball stops on this number, for a net gain of 35 dollars; otherwise, he loses his dollar bet. Find the expected value for his winnings.

Expected Value

In Las Vegas the roulette wheel has a 0 and a 00 and then the numbers 1 to 36 marked on equal slots; the wheel is spun and a ball stops randomly in one slot. When a player bets 1 dollar on a number, he receives 36 dollars if the ball stops on this number, for a net gain of 35 dollars; otherwise, he loses his dollar bet. Find the expected value for his winnings.

$$35 \cdot \frac{1}{38} + -1 \frac{37}{38} =$$

Expected Value

In Las Vegas the roulette wheel has a 0 and a 00 and then the numbers 1 to 36 marked on equal slots; the wheel is spun and a ball stops randomly in one slot. When a player bets 1 dollar on a number, he receives 36 dollars if the ball stops on this number, for a net gain of 35 dollars; otherwise, he loses his dollar bet. Find the expected value for his winnings.

$$35 \cdot \frac{1}{38} + -1 \frac{37}{38} = -0.052 \tag{1}$$

Expected Value

In a second version of roulette in Las Vegas, a player bets on red or black. Half of the numbers from 1 to 36 are red, and half are black. If a player bets a dollar on black, and if the ball stops on a black number, he gets his dollar back and another dollar. If the ball stops on a red number or on 0 or 00 he loses his dollar. Find the expected winnings for this bet.

Expected Value

In a second version of roulette in Las Vegas, a player bets on red or black. Half of the numbers from 1 to 36 are red, and half are black. If a player bets a dollar on black, and if the ball stops on a black number, he gets his dollar back and another dollar. If the ball stops on a red number or on 0 or 00 he loses his dollar. Find the expected winnings for this bet.

$$1 \cdot \frac{18}{38} - 1 \cdot \frac{20}{38} =$$

Expected Value

In a second version of roulette in Las Vegas, a player bets on red or black. Half of the numbers from 1 to 36 are red, and half are black. If a player bets a dollar on black, and if the ball stops on a black number, he gets his dollar back and another dollar. If the ball stops on a red number or on 0 or 00 he loses his dollar. Find the expected winnings for this bet.

$$1 \cdot \frac{18}{38} - 1 \cdot \frac{20}{38} = -0.052 \tag{2}$$

Is Entropy Non-negative?

We know that

$$\log p(x) \leq 0 \tag{3}$$

Is Entropy Non-negative?

We know that

$$\log p(x) \leq 0 \tag{3}$$

ff $0 \leq x \leq 1$. Thus,

$$-\log p(x) > 0. \tag{4}$$

Is Entropy Non-negative?

We know that

$$\log p(x) \leq 0 \tag{3}$$

ff $0 \leq x \leq 1$. Thus,

$$-\log p(x) > 0. \tag{4}$$

And multiplying by a non-negative probability means

$$-p(x) \log p(x) \geq 0, \tag{5}$$

so their sum is non-negative.