Probability Space

{Zamor, Okuna, Prasnik, Prath}

This is an example of a probability space $\omega \in \Omega$

Events are subsets $E \subset \Omega$

Example:

$E =$ {Zamor, Okuna}, $F =$ {Okuna, Prath}

Unions and Intersections of Events

Union: $E \cup F =$ {Zamor, Okuna, Prath}

Intersection: $E \cap F =$ {Okuna}

Probability Measures

A probability measure is a function defined on events $\mu(E)$

- $\mu(E) \geq 0$
- $\mu(\Omega) = 1$
- if $E \cap F = \emptyset$ then $\mu(E \cup F) = \mu(E) + \mu(F)$

"the probability of disjoint events is the sum of the probabilities of the events"

Example: each is equally likely (1/4)

then: μ({Zamor,Prath}) = 1/2

Mistakes

The head of a major tech corporation made a racist post on social media.

Which is more likely

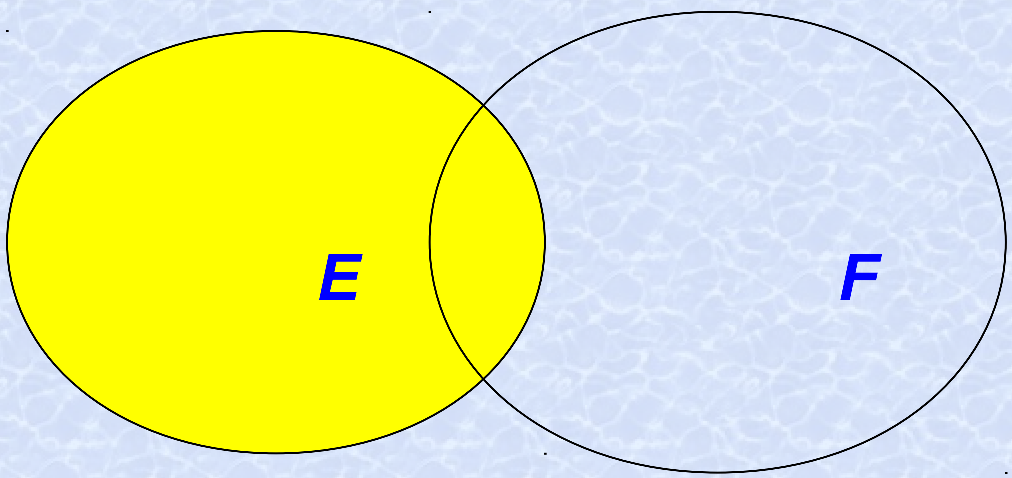
1. he is an engineer

2. he is an engineer and a white supremacist

Conditional Probability

$\mu(E|F) = \mu(E \cap F)/\mu(F)$

what does F tell you about E? information or learning



$E =$ {Prasnik}, $F =$ {begins with P} = {Prasnik, Prath}

$\mu(E) = 1/4$ but $\mu(E|F) = 1/2$

Independent Events

two events are independent if they are each uninformative about the other

$$\mu(E|F) = \mu(E)$$

From the definition of conditional probability

$$\mu(E \cap F)/\mu(F) = \mu(E)$$

or

$$\mu(E \cap F) = \mu(E)\mu(F)$$

Which is the definition of independence

Example: two coins are flipped, what is the probability both are heads?

As the coins are independent the probability of the intersection of the events the first is heads and the second is heads is the product of the probabilities: that is $(1/2)(1/2) = (1/4)$

Random Variables

A random variable $x : \Omega \rightarrow \Re$ assigns each point in the probability space a real number

For example: the random variable is income

Zamor 20K, Okuna 16K, Prasnik 40K, Prath 16K

notice that it makes perfectly good sense to add, subtract, multiply, etc. random variables

probabilities of random variables are computed from the underlying probability space

Example

x is the income

x(Prasnik) = 40

Pr(x = 16) = 1/2

Pr($2x$ = 40) = 1/4

Expectation

The expectation of a random variable is the probability weighted average

$$Ex = \sum_{\omega \in \Omega} x(\omega)\mu(\omega)$$

Example

$Ex =$ (1/4)20 + (1/4)16 + (1/4)40 + (1/4)16 = 23

Some Important Facts

$$E(x+y) = Ex + Ey$$

if *a* is a constant

$$Eax = aEx$$

but if *y* is a random variable

$Eyx \neq EyEx$ in general

if it is true the random variables are said to be uncorrelated

if the random variables are independent (events defined by the random variables are independent) then they are uncorrelated and this is true

however, random variables may fail to be independent even if uncorrelated

Variance

The variance of a random variable is

$$\text{var}x = E(x - Ex)^2 = Ex^2 - (Ex)^2$$

derivation

$$E(x - Ex)^2 = E(x^2 - 2xEx + (Ex)^2)$$

use the additive and multiplicative properties of expectation

$$= Ex^2 - 2(Ex)^2 + (Ex)^2$$

Expectation measures central tendency

variance measures uncertainty or risk

Example

var x = (1/4) $(-3)^2$ + (1/4) $(-7)^2$ + (1/4) $(17)^2$ + (1/4) $(-7)^2$ = 99

Variance and Correlation

Failure of $Exy = ExEy$

Suppose $y = x$ then this would read $Ex^2 = (Ex)^2$ meaning the variance is zero

Standard Deviation

is the square-root of the variance

measured in conformable units to the mean

it makes sense to talk about “within two standard deviations of the mean”

Example:

var$x$ = 99 sd$x$= 9.949874

rule of thumb: 95% of the time within two standard deviations of the mean

so income can be thought of as 23K plus or minus 20K

Conditional Expectation

Recall that $F =$ {begins with P}

Zamor 20K, Okuna 16K, Prasnik 40K, Prath 16K

$E(x|F) =$ 28K

Stocks and Returns

Return is x

	AAA	BBB
2020	8%	3%
2021	-2%	6%
2022	12%	3%

years are equally likely (each has probability 1/3)

$$E(x|AAA) = 6, E(x|BBB) = 4$$

so AAA give a substantially higher expected return than BBB

$$\mathrm{var}(x|AAA) = 52, \mathrm{var}(x|BBB) = 3$$

so BBB is much less risky

Large Independent Samples

repeatedly draw from the same population

for example: polling over voting, other surveys

- with replacement: independent draws, may interview people twice
- without replacement: only draw each person once
- in a large population it doesn't really matter since the chances of drawing people twice is very small
- theory based on independent draws as this is about right in a large population

Central limit theorem: the average follows approximately a "normal" distribution

In particular 95% of the time the average will be within two standard deviations of the true mean

Binomial Sampling

each voter has an independent probability p of voting for Labour

we poll a number N voters, record the random variable x equal to 1 if they say they are voting for Labour and 0 if the say they are not

mean and variance of x

$$Ex = (1-p)0 + p1 = p$$

$$\mathrm{var}x = (1-p)(0-p)^2 + p(1-p)^2 = p(1-p)$$

Sample Average

sample average: $\overline{x} = (1/N)\sum_{i=1}^{N} x_i$

mean: $E\overline{x} = (1/N)NEx_i = p$

variance: $\mathrm{var}\overline{x} = (1/N)^2 N\mathrm{var}x_i = p(1-p)/N$

standard deviation $\mathrm{sd}\overline{x} = \sqrt{p(1-p)}/\sqrt{N}$

does not depend on population size

Diversification

many risky stocks with independent returns

expected return $\overline{x}$ and variance σ^2

put 100 dollars in one stock

expected return $100x$ variance $10000\sigma^2$

put 1 dollar in each stock

expected return $100x$ variance $100\sigma^2$

diversification reduces risk

for example: investing in a index mutual fund that holds many stocks rather than in an individual stock entails less risk

What if Returns Are Not Independent?

finance experts compute β measuring how correlated the stock is with the market

adding a stock with $\beta > 1$ increases risk, $\beta < 1$ reduces it

Concepts

- **probability**, **conditional probability**
- **random variable**, **expectation**, **variance**
- **risk**, **diversification**

Skill

Given information about stock returns

Find the expected return and variance