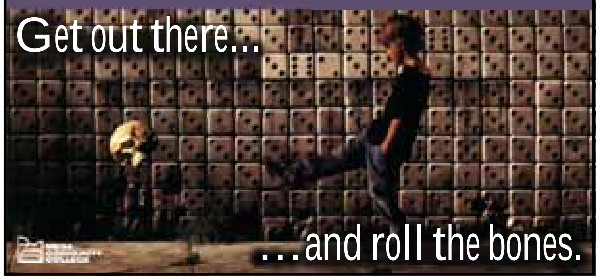
Dr. Derek Borman

PROBABILITY

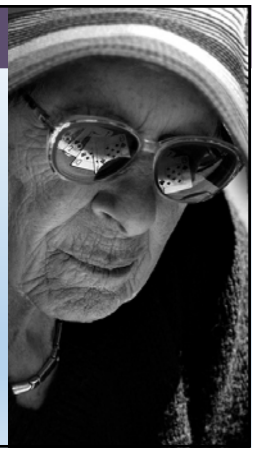
Moving from uncertainty to predictability

INTRODUCTION TO STATISTICS



LEARNING OBJECTIVES

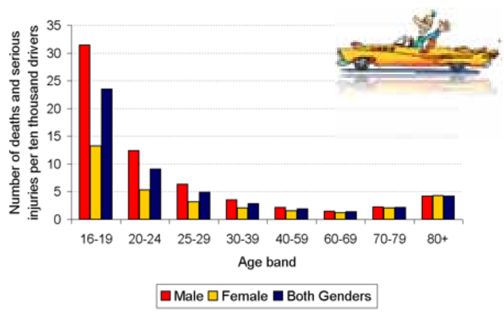
- Define and discuss probability in our everyday lives
- Understand the differences between the addition and multiplication rules
- Distinguish between independent and dependent probability events
- Understand the basics of conditional probability
- Distinguish among the different forms of subjective probability.
- Be able to perform all requisite calculations



Categories of Probability

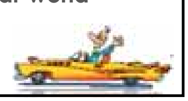
- Real World: **Based on historical/empirical assessments of chance in real life**
- Theoretical: **All possible outcomes equally likely**
- Subjective: **Probability estimates based on personal judgment**

Real World Probability and Insurance

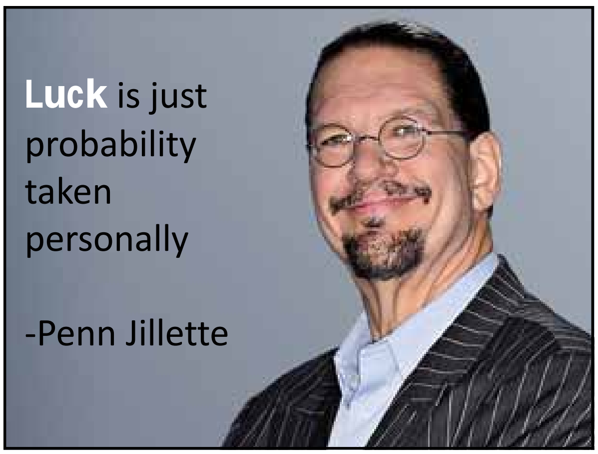


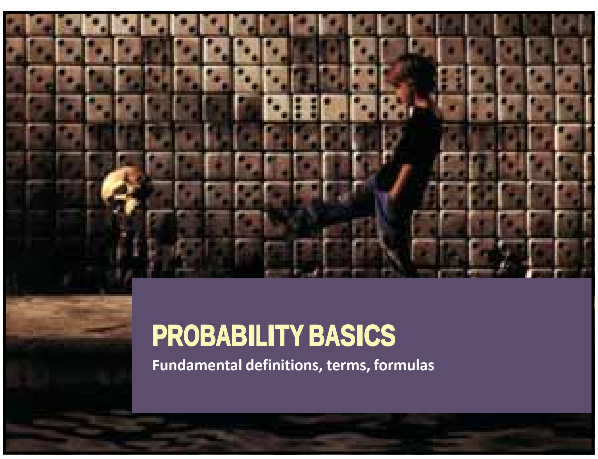
Real World Probability and Insurance

- Get into small groups and construct two profiles:
  - The most insurable driver in the world
  - The least insurable driver in the world
    - What are these people like? Demographic characteristics? Habits?
- List your characteristics and share.
- How do your profiles relate to real-world probability?



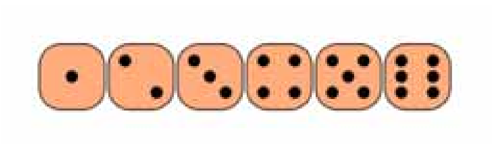

Things Said Just Prior to an Accident
Watch out!
Oh #$%@!!!!
God save us!
Hold on tight!!!
This is going to hurt
Lets make a video for YouTube!
Hold my beer and watch this!!!!!!
GraphJam.com


Luck is just probability taken personally
-Penn Jillette


PROBABILITY BASICS
Fundamental definitions, terms, formulas

Definitions: Space & Events

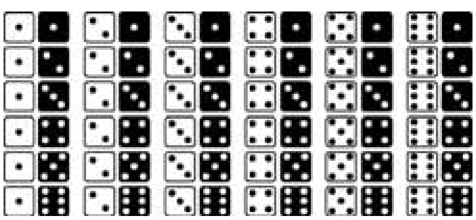
- Sample space: **all possible outcomes**
  - **What is the sample space for a single role of the die?**



- Event: **favorable outcome within a sample space**

Definitions: Space & Events

- **How about the sample space for two roles of the dice?** Exponential increase in possible outcomes



Probability Definition

- The probability of an event is equal to the number of ways a desired outcome can occur divided by all possible outcomes--over the long run.

$$\text{Probability} = \frac{\text{\# favorable outcomes}}{\text{\# of possible, equally-likely outcomes}}$$

OR (Addition) Vs. AND

- OR: Probability of obtaining one OR another outcome equals sum of their individual probabilities

$$p(\text{A or B}) = p(\text{A}) + p(\text{B})$$

What is the probability of rolling a 6 OR a 1 on a single role of a single die? A 2, 3, or 5?

See slide notes for probability when rolling two dice as in craps.

AND (Multiplication) Vs. OR

- AND: probability of obtaining one AND another outcome successively equals product of their probabilities
  - MULTIPLE DESIRABLE EVENTS IN A SERIES

$$p(\text{A, B}) = p(\text{A}) \times p(\text{B})$$

What is the probability of rolling a 1 AND a 1 on two successive rolls of the dice? How about a 2, 3, AND 5?

Probability Properties

The probability of an event, say event A, is denoted P(A)

Conditional probability denoted P(B|A), read as "probability of B given occurrence of A"

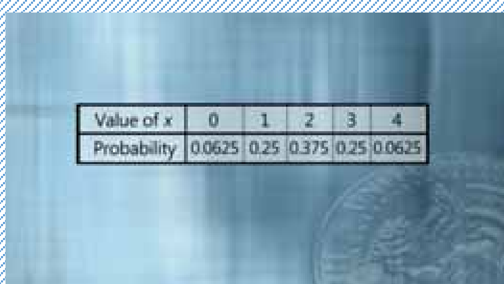
All probabilities are between 0 and 1 (i.e. $0 < P(A) < 1$)

Sum all probabilities in sample space = 1.0 or 100%

Independent Vs. Dependent

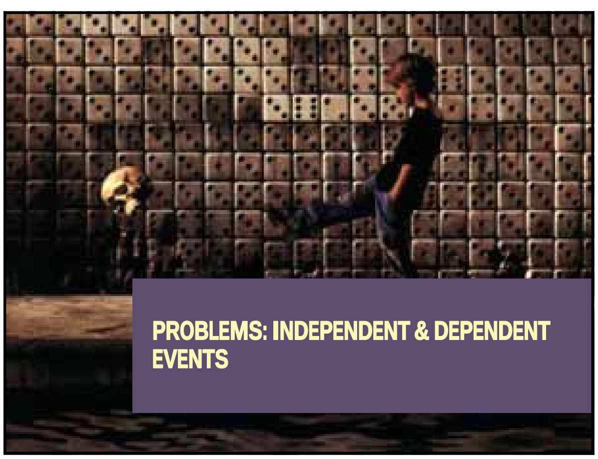
- **Independent events**: Occurrence of one event does not affect probability of later events
  - Events independent when: **P(A | B) = P(A) and P(B | A) = P(B)**
    - Replace card after each pull from deck
- **Dependent events**: occurrence of one event changes the probability of another events
  - Events dependent when: **P(B | A) ≠ P(B)**
    - Draw cards from a deck and DON'T replace after each pull

Challenger Disaster (1986)



Value of x	0	1	2	3	4
Probability	0.0625	0.25	0.375	0.25	0.0625

Discerning independent and dependent events is not an easy task…



PROBLEMS: INDEPENDENT & DEPENDENT EVENTS

How likely are certain "poker" draws if you put the card back after each draw?

- One pair (e.g., 10-10)
- Two pair (e.g., K-K-9-9)
- Three of a kind (e.g., 4-4-4)
- Straight (e.g., 9-8-7-6-5)
- Full house (e.g., 2-2-2-Q-Q)
- **Four of a kind (e.g., J-J-J-J)**
- Straight flush (e.g., Q-J-10-9-8)

$$p(A, B) = p(A) \times p(B)$$

How likely are certain "poker" draws if you put the card back after each draw?

J-J-J-J

4/52 x 4/52 x 4/52 x 4/52

.08 x .08 x .08 x .08 = .00004

There is a four in one-hundred thousand chance of doing this

Remember that .00004 is not a percentage. In order to turn this fraction to a %, we must multiply by 100.

Conditional Probability

- Conditional probability: **process of assessing probability among dependent events**
  - **Probability of event B given event A:**

$$p(A, B) = p(A) \times p(B \mid A)$$

- One event affects probability of later event
- Can be a way of adding information to improve accuracy of estimates

How likely are certain "poker" draws if you don't replace the card each time?

- One pair (e.g., 10-10)
- Two pair (e.g., K-K-9-9)
- Three of a kind (e.g., 4-4-4)
- Straight (e.g., 9-8-7-6-5)
- Full house (e.g., 2-2-2-Q-Q)
- Four of a kind (e.g., J-J-J-J)
- Straight flush (e.g., Q-J-10-9-8)

p(A, B, C, D) = p(A) x p(B | A) x p(C | A, B) x p(D | A, B, C)

How likely are certain "poker" draws if you don't replace the card each time?

J-J-J-J

4/52 x 3/51 x 2/50 x 1/49

.08 x .06 x .04 x .02 = .000004

p(A, B, C, D) = p(A) x p(B | A) x p(C | A, B) x p(D | A, B, C)

Now, what about J-10-9-8, in THAT order?

How likely are certain "poker" draws if you don't replace the card each time?

J-10-9-8 (In that order)

4/52 x 4/51 x 4/50 x 4/49

.08 x .078 x .08 x .08 = .00004

p(A, B, C, D) = p(A) x p(B | A) x p(C | A, B) x p(D | A, B, C)

Now, what about J-10-9-8, in ANY order?

How likely are certain "poker" draws if you don't replace the card each time?

J-10-9-8 (In any order)

16/52 x 12/51 x 8/50 x 4/49

.31 x .23 x .16 x .08 = .0009

p(A, B, C, D) = p(A) x p(B | A) x p(C | A, B) x p(D | A, B, C)

Mutually Exclusive vs. Overlapping

- Mutually exclusive events (disjoint): **not possible for both occur same time**
  - ***Events disjoint when: p*(A or B) = *p*(A) + *p*(B)**

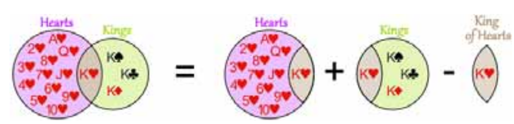
Can't roll one AND six on same roll; can't get a King and a 10 on single card pull

Overlapping Vs. Mutually Exclusive

- Overlapping or non-exclusive: **possible for events occur same time (events have overlapping properties)**
- ***Events are overlapping when:*** $p(\text{A or B}) \neq p(\text{A}) + p(\text{B})$

Can rain in Mesa and Fountain Hills; can get a card that is both an Ace and a Spade

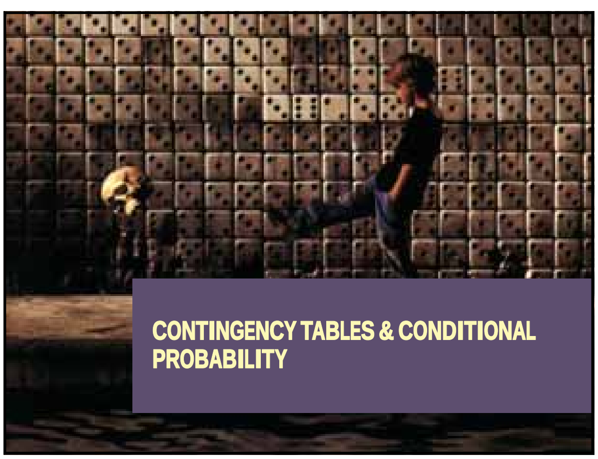
Overlapping Events



p(A or B) = p(A) + p(B) - p(A and B)

p(Ace) + p(Spade)- p(Ace and Spade)

p = .077 + .25 - .02 = .307



CONTINGENCY TABLES & CONDITIONAL PROBABILITY

Contingency Tables

- Contingency tables: frequency or percent statistics organized for two or more categorical variables
  - comparing AND with GIVEN values yields rough indication of independence/dependence between variables
  - conditional assessments can improve prediction accuracy

Write down the following data pertaining to students in ASU's College of Arts & Sciences:

53% of students are female
7% are Psychology majors
6% are female Psychology majors

<<Fill in contingency table on next slide>>

Contingency Tables

	GENDER		
MAJOR	Male	Female	Total
Psychology	**1**	**8**	**9**
Other	**62**	**63**	**125**
Total	**63**	**71**	**134**

Contingency Tables

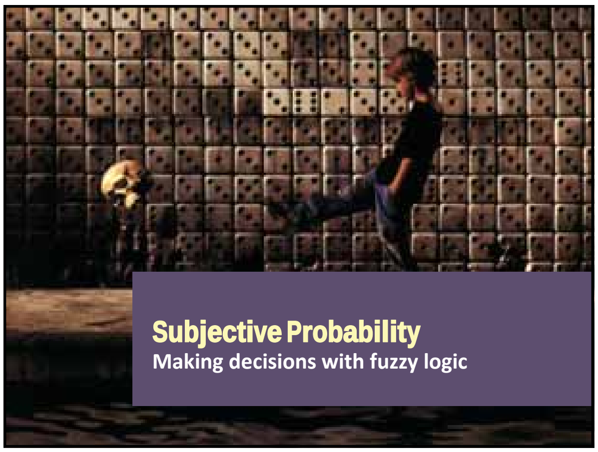
	GENDER		
MAJOR	Male	Female	Total
Psychology	.01	.06	.07
Other	.46	.47	.93
Total	.47	.53	1.00

- Column and row totals
  - Add across the columns & down the rows
- Grand total
  - Add column/row totals down and across
  - Grand total always equals 1.00 or 100%

Contingency Tables

	GENDER		
MAJOR	Male	Female	Total
Psychology	.01	.06	.07
Other	.46	.47	.93
Total	.47	.53	1.00

1. What is the probability of being female **AND** a psychology student?
2. Now, **GIVEN** that a person is majoring in psychology, what is the probability that the person is female?
3. Based on your two answers above, would you include that gender and choice of major are independent or dependent?
4. How did our **GIVEN** condition improve the accuracy of making predictions about gender and choice of psychology major?



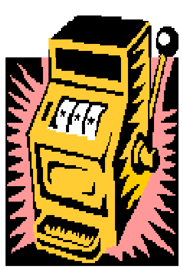
Subjective Probability

Making decisions with fuzzy logic

The Gambler's Fallacy

C'MON…I'M DUE!!

- Why do we keep pulling the lever?
  - The fallacy of the addicted gambler is the belief that the probability of a single win increases after successive losses.
- Where else do we see this kind of thinking?



AVAILABILITY HEURISTIC

If someone in the U.S. dies between the ages of 25 and 44, what is the probability of dying from…

1. a motor vehicle accident?
2. cancer?
3. suicide?
4. homicide?
5. heart disease?

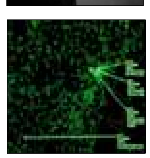
AVAILABILITY HEURISTIC

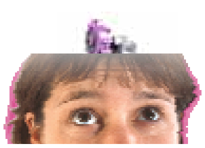
- In all cases, the real probability is less than a fraction of a percent. Why do we overestimate?
  - People tend to overestimate the probability of events based on how easily these events come to mind.

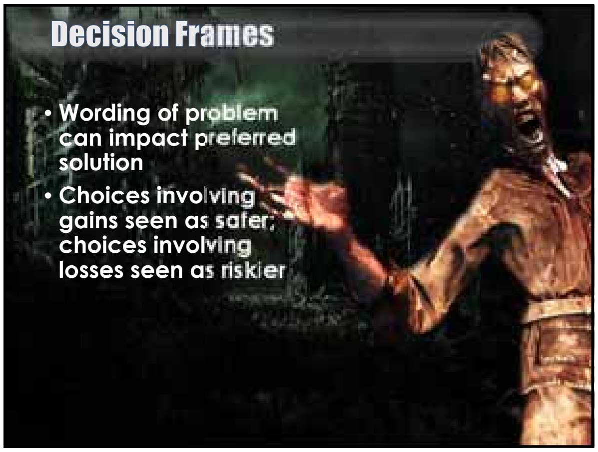
Causes of Death (age 25-44 years)	Number of Deaths	Deaths per 100,000	Probability
Motor vehicle accidents	14,528	17.3	.0002 / .02%
Cancer	22,147	26.4	.0003 / .03%
Suicide	12,536	15	.0002 / .02%
Homicide and legal intervention	9,261	11.1	.0001 / .01%
Heart disease	3,418	4.1	.00004 / .004%

Availability

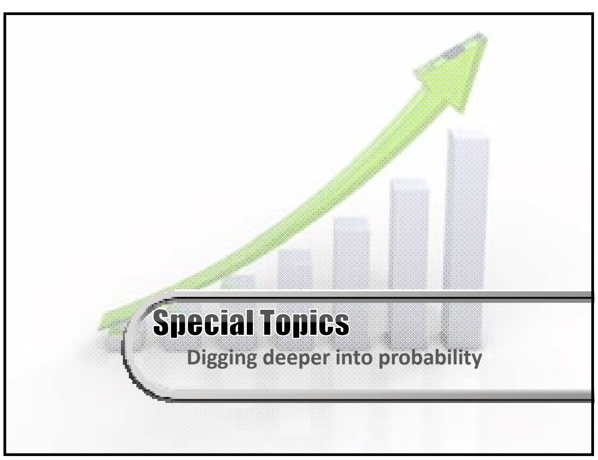
- How many planes were brought down by terrorists on 9/11/01? How many planes were up in U.S. airspace during the attacks?
- People tend to overestimate the probability of events based on how easily these events come to mind.






Decision Frames
• Wording of problem can impact preferred solution
• Choices involving gains seen as safer; choices involving losses seen as riskier


Zombie Apocalypse
PROBLEM: Imagine that the world is preparing for a zombie apocalypse, which is expected to zombify a lot of people. Two alternative programs to combat the hordes of walking dead have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:
1. If Program A is adopted, there is a 25% chance that people will be saved.
2. If Program B is adopted, there is a 75% chance that people will not be saved.
Which program would people favor most? Why?

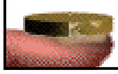

Special Topics
Digging deeper into probability

Probability: A Function Of An Observer's Knowledge

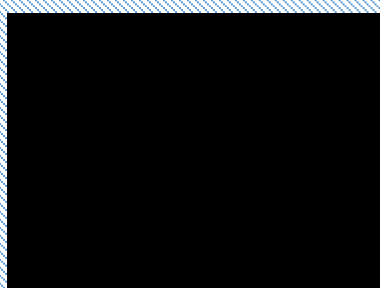
Does probability even exist?

- Binomial 50/50
  - Coin tosses over the short run…maybe 60/40
  - Coin tosses over the long run…probably 50/50

But what if we practice our flipping? Could someone flip 60, 70 or even 80% heads over 10 tosses? Give it a try!



Percy's Thoughts On The Matter



Extra Credit: Let's Make A Deal!

- Game show setting. There are 3 doors, behind one of which is a prize. Monty Hall, the host, asks you to pick a door, any door. You pick door A. Monty opens door B and shows that there is nothing behind door B. You are given the choice to stay with A or switch to door B. SHOULD YOU SWITCH? WHY?

Extra Credit: Let's Make A Deal!

- (2 EC points) Go out and read up on the "Monty Hall" probability dilemma. A quick Google search with terms like "probability and Monty Hall" or "probability and Let's Make a Deal" will yield a lot of hits. Read until you understand it.
  - Now, answer the first two questions posed: 1) Should you switch? 2) Why? Your responses must be in your own words. Don't just copy the thoughts of someone else. On a sheet of NOTEBOOK PAPER, write up your answers to these questions and attach a PRINTOUT of the website that you used. Staple and submit both.

You say it's your Birthday???

- What are the odds that somebody else in the same room has the same birthday as you?

Group of 5 people....

1. Let A = event no one in group shares same birthday
2. Then A^c = event at least 2 people share same birthday

$$P(A) = 365/365 \times 364/365 \times 363/365 \times 362/365 \times 361/365 = 0.973$$

$$P(A^C) = 1 - 0.973 = 0.027$$

That is, about a 3% chance that in a group of 5 people at least two people share the same birthday.

Group of 23 people....

Let A = event no one in group shares same birthday.

Then A^c = event at least 2 people share same birthday.

$P(A) = 365/365 \times 364/365 \times \ldots \times 343/365 = 0.493$

$P(A^C) = 1 - 0.493 = 0.507$

That is, about a 50% chance that in a group of 23 people at least two people share the same birthday.

Group of 50 people....

Let A = event no one in group shares same birthday.

Then A^c = event at least 2 people share same birthday.

$P(A) = 365/365 \times 364/365 \times \ldots \times 316/365 = 0.03$

$P(A^C) = 1 - 0.03 = 0.97$

It is a virtual certainty that in a group of 50 people at least two people share the same birthday.