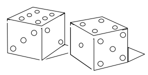
III. Probability

PBAF 527
Winter 2005



3 - 1

Learning Objectives

1. Define Experiment, Outcome, Event, Sample Space, & Probability
2. Explain How to Assign Probabilities
3. Use Venn Diagrams & Contingency Tables
4. Describe & Use Probability Rules
5. Bayes Theorem
6. Learn to Draw a Random Sample
7. Welfare Case--Probabilities

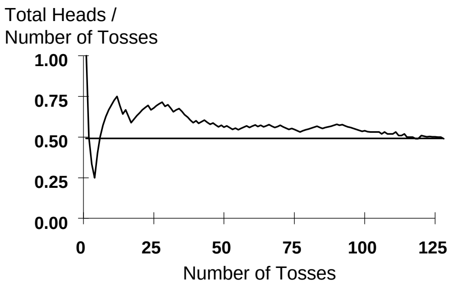
3 - 2

What's the probability of getting a **head** on the toss of a single fair coin? Use a scale from **0** (**no way**) to **1** (**sure thing**).

So toss a coin twice. Do it! Did you get one head & one tail? What's it all mean?

3 - 3

Many Repetitions!



3 - 4

What is probability?

A quantitative measure of uncertainty.

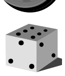
Knowing the past performance of a population (or a series of coin tosses) allows us to make statements about samples that are drawn.

3 - 5

Classical Probability

- Based on the idea that certain occurrences are equally likely to happen.

Heads or tails of a flipped coin $p(E) = \frac{1}{2}$



Roll of a die $p(E) = \frac{1}{6}$

3 - 6

Some Definitions

1. Experiment
   - Process of Obtaining an Observation, Outcome or Simple Event
2. Sample Point
   - Most Basic Outcome of an Experiment
3. Sample Space (S)
   - Collection of *All* Possible Outcomes
4. Event
   - Subset the sample space, a set of basic outcomes

Sample Space Depends on Experimenter!

3 - 7

Sample Space

Experiment	Sample Space
Toss a Coin, Note Face	Head, Tail
Toss 2 Coins, Note Faces	HH, HT, TH, TT
Select 1 Card, Note Kind	2♥, 2♦, ..., A♠ (52)
Select 1 Card, Note Color	Red, Black
Play a Football Game	Win, Lose, Tie
Inspect a Part, Note Quality	Defective, OK
Observe Gender	Male, Female

3 - 8

Outcome Properties

1. Mutually Exclusive
   - 2 Outcomes Can Not Occur at the Same Time
     - **Both** Male & Female in Same Person
2. Collectively Exhaustive
   - 1 Outcome in Sample Space Must Occur
     - Male or Female

Experiment: Observe Gender

3 - 9

Events

1. A set of basic outcomes
2. Simple Event
   - Outcome With 1 Characteristic
3. Compound Event
   - Collection of Outcomes or Simple Events
   - 2 or More Characteristics
   - Joint Event Is a Special Case
     - 2 Events Occurring Simultaneously

3 - 10

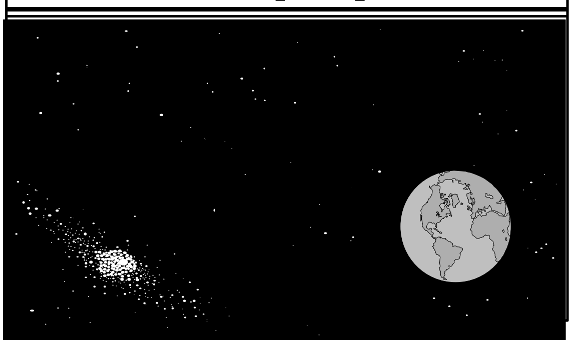
Event Examples

Experiment: Toss 2 Coins. Note Faces.

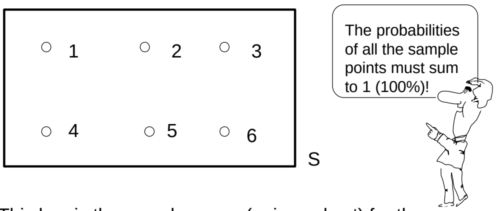
Event	Outcomes in Event
Sample Space	HH, HT, TH, TT
1 Head & 1 Tail	HT, TH
Heads on 1st Coin	HH, HT
At Least 1 Head	HH, HT, TH
Heads on Both	HH

3 - 11

Visualizing Sample Space



Venn Diagrams

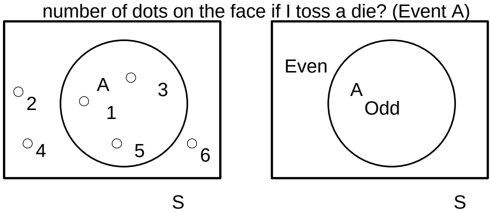


This box is the sample space (universal set) for the experiment where we toss a die. All possible sample points are in it.

3 - 13

Venn Diagram Variations

- Often we want to know about a compound event – some combination of sample points.
- What's the probability of ending up with an odd number of dots on the face if I toss a die? (Event A)



3 - 14

Probabilities are like relative frequencies

Probability of an outcome is the proportion of times it would occur in the long run (in repeated samples).

Probability of A (an event) is a numerical measure of the likelihood of the event A happening:

$$P(A) = \frac{n(A)}{n(S)}$$

where n(A)= the number of elements in the set of the event A
where n(S)=the number of elements in the sample space

3 - 15

So, in practice…

The **probability** is the proportion of that event in the sample space.

Sometimes, you are given those proportions.

Sometimes, you are only given counts and need to calculate the probabilities.

3 - 16

Learning Objectives

1. Define Experiment, Outcome, Event, Sample Space, & Probability
2. Explain How to Assign Probabilities
3. Use Venn Diagrams & Contingency Tables
4. Describe & Use Probability Rules
5. Bayes Theorem
6. Learn to Draw a Random Sample
7. Welfare Case--Probabilities

3 - 17

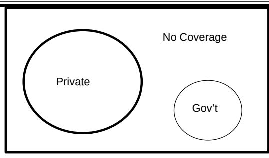
Example: Insurance Coverage in the U.S.

A recent study of US adult population examined the relationship between medical insurance coverage and doctor visits over a 24-month period.

How does coverage affect the probability of a doctor visit?

3 - 18

Insurance Coverage in the U.S.



Proportion of Cases with Each Combination of Coverage and Visits:

	Doctor Visits		
Medical Coverage	**One or More**	**None**	**Total**
Private	.72	.13	**.85**
Government	.09	.01	**.10**
None	.03	.02	**.05**
Total	**.84**	**.16**	**1.0**

Simple Event → Government

3 - 19

Using a Contingency Table

What is the probability that a person in the sample had

- medical insurance?
- at least one doctor visit?
- insurance and visited the doctor?

Proportion of Cases with Each Combination of Coverage and Visits:

	Doctor Visits		
Medical Coverage	**One or More**	**None**	**Total**
Private	.72	.13	**.85**
Government	.09	.01	**.10**
None	.03	.02	**.05**
Total	**.84**	**.16**	**1.0**

3 - 20

Remember…

Probabilities for an "outcome set" (sample space) are:

- Mutually exclusive
- Exhaustive (have to be in one category)
- Whole group must sum to 1.
  All probabilities are between 0 and 1
  (0=$P(A)$ =1)

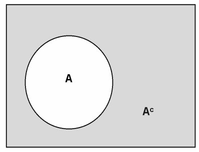
You can calculate the total probability of an event by adding up mutually exclusive categories.

3 - 21

Complement

A **set** is a collection of elements

- Empty set: no elements, ϕ
- Universal set: all elements, S—the sample space



The **complement** of set A (A^c) is the set of all elements in the universal set that are not in the set A.

Rule of Complements $P(A^c) = 1 - P(A)$

3 - 22

Using a Contingency Table (2)

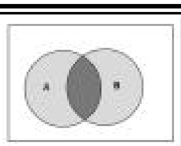
	Doctor Visits		
Medical Coverage	One or More	None	Total
Private	.72	.13	**.85**
Government	.09	.01	**.10**
None	.03	.02	**.05**
Total	**.84**	**.16**	**1.0**

What is the probability that a person in the sample

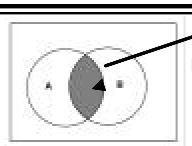
- Will have no medical insurance?
- Did not see a doctor during this two-year period?

3 - 23

Union and Intersection



A U B



A n B

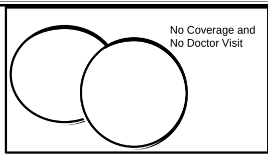
The **union** of A and B is the set containing all elements in A or B or both. Shown as U

The **intersection** A and B is the set containing all elements of A and B. Shown as n The joint probability

Rule of Unions ("or" events) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

3 - 24

Event Union



What's the probability of either visiting the doctor or having insurance?

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

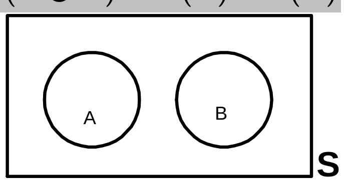
	Doctor Visits		
Medical Coverage	One or More	None	Total
Private	.72	.13	.85
Government	.09	.01	.10
None	.03	.02	.05
Total	.84	.16	1.0

3 - 25

For Mutually Exclusive Events

$$P(A \cap B) = 0$$

$$P(A \cup B) = P(A) + P(B)$$



3 - 26

Event Probability Using Contingency Table

	Event		
Event	[illegible]	[illegible]	[illegible]
A_1	[illegible]	[illegible]	[illegible]
A_2	[illegible]	[illegible]	$P(A_2)$
Total	$P(B_1)$	$P(B_2)$	1

Joint Probability Marginal (Simple) Probability

3 - 27

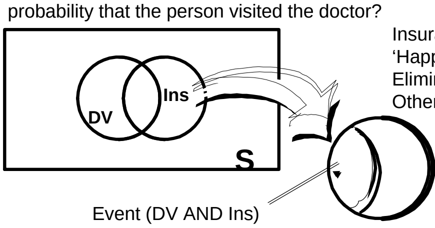
Conditional Probability

1. Event Probability ***Given*** that Another Event Occurred
2. Revise Original Sample Space to Account for ***New*** Information
   - Eliminates Certain Outcomes
3. P(A **|** **B**) = P(A n B) / P(**B**)

3 - 28

Conditional Probability Using Venn Diagram

Given that a person in our sample has insurance, what is the probability that the person visited the doctor?



Insurance 'Happens': Eliminates All Other Outcomes

Event (DV AND Ins)

Insurance is the revised Sample Space

3 - 29

Conditional Probability Using Contingency Table

Given that a person in our sample has insurance, what is the probability that the person visited the doctor?

	Doctor Visits		
Medical Coverage	One or More	None	Total
Private	.72	.13	.85
Government	.09	.01	.10
None	.03	.02	.05
Total	.84	.16	1.0

Revised Sample Space

P(DV|Ins) = P(DV and Ins)/P(Ins) = .81/.95 = .85

3 - 30

Conditional Probability Using Contingency Table (2)

If someone is uninsured, what is the probability that the person had at least one doctor visit?

	Doctor Visits		
Medical Coverage	**One or More**	**None**	**Total**
Private	.72	.13	**.85**
Government	.09	.01	**.10**
None	.03	.02	**.05**
Total	**.84**	**.16**	**1.0**

Revised Sample Space

P(DV|Ins) = P(DV and Ins)/P(Ins) = .03/.05 = .60

3 - 31

Multiplicative Rule

1. Used to Get Compound Probabilities for ***Intersection*** of Events
   - Called Joint Events
2. P(A and B) = P(A ∩ B)
   = P(**A**)*P(B|**A**)
   = P(**B**)*P(A|**B**)

Sometimes we have one probability and want to find another

3 - 32

Multiplicative Rule Example

What is the probability of someone in this sample having insurance and no doctor visits?

	Doctor Visits		
Medical Coverage	**One or More**	**None**	**Total**
Private	.72	.13	**.85**
Government	.09	.01	**.10**
None	.03	.02	**.05**
Total	**.84**	**.16**	**1.0**

P(A n B) = P(**A**)*P(B|**A**)
= P(**B**)*P(A|**B**)

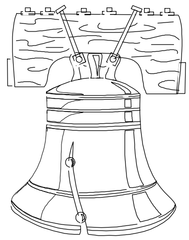
P(Ins and No DV)
= P(no DV | **Ins**)P(**Ins**)
= P(Ins| **no DV**)P(**no DV**)

(Last example, P(DV|Ins)=.85, so we can use that)
=[1-P(DV|Ins)]P(Ins) =[1-.85].95=.14

3 - 33

Statistical Independence

1. When Event Occurrence Does ***Not*** Affect Probability of Another Event
   - Toss 1 Coin Twice
2. Causality Not Implied
3. Tests For
   - P(**A** | B) = P(**A**)
   - P(**B** | A)= P(**B**)
   - P(A and B) = P(A)*P(B) ← Product Rule



Events that are not independent are dependent.

3 - 34

Independence Example

Is the chance of visiting a doctor statistically independent of having insurance?

If independent:

$$P(DV|INS) = P(DV)$$
$$P(INS|DV) = P(INS)$$
$$P(DV \cap INS) = P(DV)P(INS)$$

If any one of these is not true, then the chance of visiting a doctor is not statistically independent of having insurance. In English?

$$P(DV|INS) = .85$$
$$P(DV) = .84$$

SO: $P(DV|INS) \neq P(DV)$

So, we don't need to check the other.

3 - 35

Independence Example (2)

	Doctor Visits		
Medical Coverage	**One or More**	**None**	**Total**
Private	.72	.13	**.85**
Government	.09	.01	**.10**
None	.03	.02	**.05**
Total	**.84**	**.16**	**1.0**

ALSO:

$$P(DV \cap INS) = .81$$
$$P(DV)P(INS) = (.95)(.84) = .798$$

SO:

$$P(DV \cap INS) \neq P(DV)P(INS)$$

3 - 36

A Product Rule for Independent Events

Probability of Intersection of Independent Events

- Probability that all events will happen
- The product of the separate probabilities

3 - 37

A skilled Evans School graduate is applying for her dream job. Even in this job market, seven public and nonprofit organizations are considering her.

- At three of the seven she is a finalist, which means that at each organization she is in the final group of three applicants, one of whom will be chosen for the position.
- At two of the seven, she is a semifinalist, that is, one of six candidates.
- At the last two, she is at an early stage of her application and believes there is a pool of about 20 candidates for each position.

Assuming that there is no exchange of information, or influence, across organizations as to their hiring decisions, and that this bright graduate is as likely to be chosen as any other applicant, what is the her probability of getting a job offer from all the places to which she applied?

3 - 38

What is the her probability of getting a job offer from all the places to which she applied?

Organization	Probability of a job offer
1	
2	
3	
4	
5	
6	
7	

3 - 39

Key Points to Remember

- Are mutually exclusive events the same as independent events?
- What is the probability that two mutually exclusive events occur?
- What is the probability that two independent events both occur?

3 - 40

Product Rule Example

A certain lie detector shows a positive reading 10% of the time when a person is telling the truth and 95% of the time when the person is lying. Suppose we have 2 suspects for a crime and only 1 is guilty.

What is the probability that the detector shows:

- A positive reading for both suspects?
- A positive reading for the guilty subject and a negative reading for the innocent—it is right?
- A positive reading for the innocent subject and a negative reading for the guilty subject—it is totally wrong?

3 - 41

Let

P1=positive reading for guilty suspect= .95

P2=positive reading for innocent suspect=.10

A positive reading for both? $P(P1 \cap P2)=$ .95 x .10=.095

A positive reading for the guilty subject and a negative reading for the innocent? $P(P1 \cap P2^c)=$ .95 x .90=.855

A positive reading for the innocent subject and a negative reading for the guilty subject—it is totally wrong?

3 - 42

Bayes Theorem

When we need to know conditional probability, but only know its opposite.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

(the denominator is simply the probability of A)

Prior Probabilities: $P(B)$ and $P(B^c)$ (We know these)

Posterior Probability: $P(B|A)$ (We calculate this)

3 - 43

HIV Screening Test
(in a given population)

- When administered to a person who has HIV, the test will be positive with probability 0.92
- When administered to a person who does not have HIV, the test will erroneously give a positive result with probability 0.04
- The rate of HIV infection in the population of interest is 0.1 percent.
  - What's the probability that a random person does not have HIV?
  - For a randomly selected person from the population who tests positive, what is the probability that they have HIV?

3 - 44

HIV Screening Test
(in a given population)

- What's the probability that a random person does not have HIV?
- For a randomly selected person from the population who tests p

$(0.92)(0.001) + (0.04)(0.999)$

3 - 45

Random Sampling

- All inferences made in statistics based on samples.
  - Want to represent population as closely as possible
  - Simple Random Sampling
    - Each element drawn with equal and known probability

3 - 46

Random Sampling

How do you select a random sample?

- Put slips of paper into a hat
- Use a random number table
- Toss a single die (if there are 6 elements in the population!).
- Random number generators
  - In Excel =Randbetween(bottom, top)
  - In SPSS, the program can select the random subsample for you.

3 - 47

Random Sampling

- A sample is just a single experiment, drawing *n* of *N* elements
  - Each different sample represents a sample point of one experiment.
- With a small population, often you can list all the samples. With a large population, it is more difficult to image all the possible samples.
- We want to make sure that each sample has an equal probability of being selected!

3 - 48

Random Sampling

To calculate the number of possible samples:

$$\binom{N}{n} = \left(\frac{N!}{n!(N-n)!} \right)$$

Combinatorial Rule

where $n! = n(n-1)(n-2)\ldots(3)(2)(1)$

3 - 49

Random Sampling

How many samples of 5 are possible from the class data?

60	62	63	63	63	64	64	65	66	66	66	67
67	68	68	69	69	69	69	69	70	71	72	

3 - 50

Random Sampling

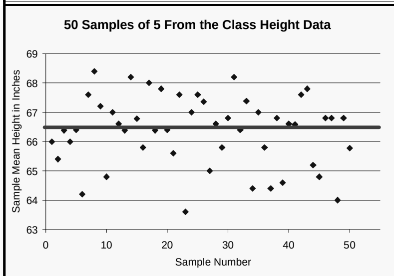
Using the random number table at the back of the book, select a random sample of 5 from the class height data.

60	62	63	63	63	64	64	65	66	66	66	67
67	68	68	69	69	69	69	69	70	71	72	

Calculate the average height for your sample

3 - 51

Random Sampling



$\bar{x} = 66.5$

5	10	20	30	40	50
66.0	66.2	66.6	66.5	66.4	66.4

3 - 52

Conclusion

1. Defined Experiment, Outcome, Event, Sample Space, & Probability
2. Explained How to Assign Probabilities
3. Used Venn Diagrams & Contingency Tables
4. Described & Used Probability Rules
5. Bayes Theorem
6. Learned to Draw a Random Sample
7. Welfare Case--Probabilities

Next time: Why does the area under the curve equal 1?

3 - 53

End of Chapter

Any blank slides that follow are blank intentionally.