111, section 7.5 Conditional Probability

notes by Tim Pilachowski

In **conditional probability** an outcome or event B is dependent upon another outcome or event A. Formally, $P(B \mid A) = P(B \text{ given } A)$ = probability that B will happen given the prior condition that A has already happened.

7.1, 7.2 & 7.4 Example D revisited: Suppose that a box contains 3 blue blocks and 2 yellow blocks.

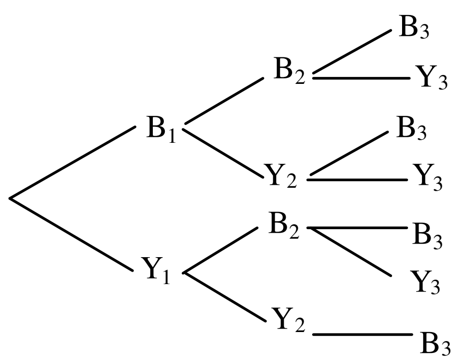
If we picked blocks and then put them back ("with replacement") the probability of picking a blue block would never change, it would always be the same. There would always be 3 blue blocks in the box, no matter what had been picked previously.

$$P(B_1) = P(B_2) = P(B_3) = \frac{\text{number of ways } B_i \text{ can happen}}{\text{number of possible outcomes}} = \frac{3}{5}$$

The picks are **independent**, like two coins tossed or two dice rolled.

The tree diagram for randomly picking three blocks ***<u>without replacement</u>***, would look like this:

Experiment: Pick three blocks ***<u>without replacement.</u>***



Let B_1 = blue on the first pick, B_2 = blue on the second pick, B_3 = blue on the third pick, Y_1 = yellow on the first pick, Y_2 = yellow on the second pick, Y_3 = yellow on the third pick. You can fill in the probabilities on the tree diagram above as we calculate them in class. Those marked with * below will be left for you to find on your own.

P(picking blue second given a blue was picked first) $= P(B_2 \mid B_1) =$

$P(Y_2 \mid B_1) =$ $\qquad$ $P(B_2 \mid Y_1) =$ $\qquad$ $P(Y_2 \mid Y_1) =$

*P(picking blue third given a blue was picked first and second) $= P(B_3 \mid B_1 \cap B_2) =$

$P(Y_3 \mid B_1 \cap B_2) =$ $\qquad$ $P(B_3 \mid B_1 \cap Y_2) =$ $\qquad$ * $P(Y_3 \mid B_1 \cap Y_2) =$

$P(B_3 \mid Y_1 \cap B_2) =$ $\qquad$ * $P(Y_3 \mid Y_1 \cap B_2) =$ $\qquad$ * $P(B_3 \mid Y_1 \cap Y_2) =$

* $P(Y_3 \mid Y_1 \cap Y_2) =$

In an "intersection", i.e. an "and" situation, moving left-to-right on the tree diagram, multiply probabilities.

P(blue first and blue second and yellow third) = * $P(B_1 \cap B_2 \cap B_3) = P(B_1) * P(B_2 \mid B_1) * P(B_3 \mid B_1 \cap B_2)$ =

$P(B_1 \cap B_2 \cap Y_3) =$

$P(B_1 \cap Y_2 \cap B_3) =$

* $P(B_1 \cap Y_2 \cap Y_3) =$

$P(Y_1 \cap B_2 \cap B_3) =$

* $P(Y_1 \cap B_2 \cap Y_3) =$

* $P(Y_1 \cap Y_2 \cap B_3) =$

In a "union", i.e. an "or" situation, moving up-and-down on the tree diagram, add probabilities.

*P(3 blue blocks) = $P(B_1 \cap B_2 \cap B_3) =$

P(2 blue blocks and 1 yellow block) =

*P(1 blue block and 2 yellow blocks) =

*P(3 yellow blocks) =

$P(S)$ =

observations about Lecture 7.4 Example D compared to Lecture 7.5 Example D:

For Example D, we were easily able to count how many blocks were left. In other situations, we won't be able to count so easily. So we have a formal definition and formula.

Earlier, we used the multiplication principle.

$$P(E) * P(F \mid E) = P(E \cap F)$$

With algebraic manipulation, we get a formula for conditional probability.

$$P(F \mid E) = \frac{P(E \cap F)}{P(E)}$$

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)} = \frac{\frac{\text{number of events in}(E \cap F)}{\text{number of events in}(S)}}{\frac{\text{number of events in}(F)}{\text{number of events in}(S)}} = \frac{\text{number of events in }(E \cap F)}{\text{number of events in }(F)}$$

In words, when we're considering only event F, the conditional probability is that portion/fraction that also includes event E ?

Example A: Given two events C and D in a sample space S, if we know that $P(C) = 0.3$ and $P(D \mid C) = 0.2$, then what is $P(C \cap D)$?

Example B: Given two events E and F in a sample space S, if we know that $P(E) = \frac{3}{4}$, $P(F) = \frac{1}{3}$ and $P(E \cap F) = \frac{1}{4}$, then a) what is $P(E \mid F)$? b) what is $P(F \mid E)$?

7.1, 7.2 & 7.4 Example D revisited: Suppose that a box contains 3 blue blocks and 2 yellow blocks.

$P(B_2 \mid B_1) =$

$P(B_2 \mid Y_1) =$

Notice the mathematical relationship. Formally, if two events E and F are independent then

$$(1)\ P(E \mid F) = P(E) \quad \text{and} \quad (2)\ P(F \mid E) = P(F).$$

Using the multiplication principle, $P(E \cap F) = P(E) * P(F \mid E)$. Then, if the two events are independent,

$$(3)\ P(E \cap F) = P(E) * P(F).$$

Any one of these three formulas is sufficient to prove or disprove independence of events.

Example A revisited: Given two events C and D in a sample space S, if we know that $P(C) = 0.3$, $P(D) = 0.4$, and $P(C \cup D) = 0.64$, are events C and D independent?

Example B revisited: Given two events E and F in a sample space S, if we know that $P(E) = \frac{3}{4}$, $P(F) = \frac{1}{3}$ and $P(E \cap F) = \frac{1}{4}$, then are events E and F independent?

Example C: For two events G and H in a sample space S, we know that they are independent and that $P(G) = 0.7$ and $P(H) = 0.4$.

$P(G \cap H) =$

$P(G \mid H) =$

$P(H \mid G^c) =$

$P(H^c \mid G^c) =$

Example E: You toss two standard six-sided dice.

$S = \{$	(1, 1),	(2, 1),	(3, 1),	(4, 1),	(5, 1),	(6, 1),
	(1, 2),	(2, 2),	(3, 2),	(4, 2),	(5, 2),	(6, 2),
	(1, 3),	(2, 3),	(3, 3),	(4, 3),	(5, 3),	(6, 3),
	(1, 4),	(2, 4),	(3, 4),	(4, 4),	(5, 4),	(6, 4),
	(1, 5),	(2, 5),	(3, 5),	(4, 5),	(5, 5),	(6, 5),
	(1, 6),	(2, 6),	(3, 6),	(4, 6),	(5, 6),	(6, 6) $\}$

A = at least one of the dice is a 4 $\quad P(A) =$

B = the sum of the two dice is 9 $\quad P(B) =$

$A \cap B =$

$P(A \cap B) =$

$A \mid B =$

$P(A \mid B) =$

$B \mid A =$

$P(B \mid A) =$

Are events A and B independent? How do you know?

Example F-1: You pick a card from a standard deck of 52. $n(S) = 52$.
S = { A-S, A-H, A-D, A-C, K-S, K-H, K-D, K-C, Q-S, Q-H, Q-D, Q-C, … , 2-S, 2-H, 2-D, 2-C }.

C = the card is an Ace $P(C) =$

D = the card is a Spade $P(D) =$

$C \cap D =$

$P(C \cap D) =$

$C \mid D =$

$P(C \mid D) =$

$D \mid C =$

$P(D \mid C) =$

Are events C and D independent? How do you know?

Example F-2: You pick two cards from a standard deck of 52.

$n(S) =$

E = at least one card is an Ace $P(E) =$

$E^c =$ $P(E^c) =$

F = picking a pair $P(F) =$

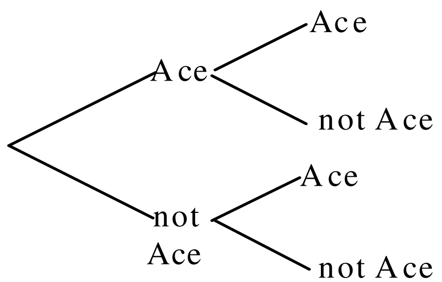
$E \cap F =$

$P(E \cap F) =$

Are events E and F independent? How do you know?

Example F-2 revisited: You pick two cards from a standard deck of 52. $n(S) = C(52, 2)$

E = at least one card is an Ace, F = picking a pair, $E \cap F$ = picking a pair of Aces



P(Ace on first pick) =

P(not-Ace on first pick) =

P(Ace | Ace) =

P(not-Ace | Ace) =

P(Ace | not-Ace) =

P (not-Ace | not-Ace) =

P(two Aces) = $P(E \cap F)$ =

P(Ace then not-Ace) =

P(not-Ace then Ace) =

P (neither card is an Ace) =

P (E) =

P (F) =

Example G: Silver Springs, Florida, has a snack bar and a gift shop. The management observes 100 visitors, and counts 55 who make a purchase in the gift shop (event G), 65 who eat in the snack bar (event H), and 40 who do both.

$G | H$ =

P ($G | H$) =

interpretation of $G | H$:

$H | G$ =

P ($H | G$) =

interpretation of $H | G$:

Example H: At Matriarch University (U. Mama), 35% of the students have an academic scholarships. Of the students who have an academic scholarship, 80% live on campus.

J = has an academic scholarship, K = lives on campus

verbal description, then event-probability notation:

0.35 =

verbal description, then event-probability notation:

0.80 =

$J \cap K =$

$P(J \cap K) = ?$

Example I: Previous quality control tests indicate that 0.5% of the spark plugs produced in a factory are defective. A case contains 4 dozen spark plugs.

a) What is the probability that a single spark plug is not defective?

b) What is the probability that a customer receives a case with at least one defective spark plug?

c) What is the probability that a customer who opens cases at random will open four cases before finding one that has no defective spark plugs?

Example J: The Gallup organization conducted almost 150,000 interviews from January through May 2009.

	Democrat	Independent	Other	Republican
Female	21%	13%	4%	13%
Male	16%	17%	2%	14%

http://www.gallup.com/poll/120839/Women-Likely-Democrats-Regardless-Age.aspx
Results are based on telephone interviews with 149,192 national adults, aged 18 and older, conducted Jan. 2-May 31, 2009, as part of Gallup Poll Daily tracking. Interviews are conducted with respondents on land-line telephones (for respondents with a land-line telephone) and cellular phones (for respondents who are cell-phone only). Categories are arranged in alphabetical order.

This table gives probabilities of intersections! $P(\text{Male} \cap \text{Independent}) = 17\%$. $P(\text{Female} \cap \text{Other}) = 4\%$.
Let A = member of the Democratic Party and B = female gender.

$P(A) =$

$P(A^c) =$

$P(B) =$

$P(B^c) =$

$A \cap B =$

$P(A \cap B) =$

$A \cap B^c =$

$P(A \cap B^c) =$

$A \cup B =$

$P(A \cup B) =$

$A^c \cup B =$

$P(A^c \cup B) =$

$A \mid B =$

$P(A \mid B) =$

interpretation:

$B \mid A =$

$P(B \mid A) =$

interpretation:

Are A and B independent events? How do you know?

Example K: In early 2010, households were surveyed about health insurance coverage.

	Under 65 years	18-64 years	Under 18 years
Uninsured %	17.5%	21.5%	7.4%
No. Uninsured (millions)	46.4	40.9	5.5
Public %	21.2%	14.4%	38.4%
Private %	62.7%	65.5%	55.5%

Data from Table 1, Table 2 and Table 3, *Health Insurance Coverage: Early Release of Estimates From the National Health Interview Survey, January–March 2010* by Robin A. Cohen, Ph.D., and Michael E. Martinez, M.P.H., M.H.S.A., Division of Health Interview Statistics, National Center for Health Statistics. **IMPORTANT:** A footnote states, "A small number of persons were covered by both public and private plans and were included in both categories." As a result, note that the percents in each column add up to greater than 100%.

This table gives conditional probabilities! $P(\text{Uninsured} \mid \text{18 - 64 years}) = 21.5\%$. $P(\text{Public} \mid \text{Under 65}) = 21.2\%$. Let C = under 18 without health insurance, D = under 18 having public health insurance and E = under 18 having private health insurance.

$D \cap E =$

$P(D \cap E) = ?$

interpretation:

$D \cup E =$

$P(D \cup E) =$

interpretation:

Let F = person is under 18 years and G = person is under 65 and lacks health insurance coverage.

$P(G) =$

interpretation:

$P(G \mid F) =$

interpretation:

Are events F and G independent? How do you know?

Note that $P(G \mid F)$ is not the same thing as $P(F \mid G)$!

interpretation of $P(G \mid F)$:

interpretation of $P(F \mid G)$:

Example L: U.S. Department of Labor employment reports place jobs into the five following categories:

Management, professional, and related occupations	(Mgt)
Service occupations	(Svc)
Sales and office occupations	(SO)
Natural resources, construction, and maintenance occupations	(NCM)
Production, transportation, and material moving occupations	(PTM)

In 2010, 39.3% of jobs were classified as Mgt, 14.5% were Svc, 23.2% were SO, 9.9% were NCM, and 13.1% were PTM. Among Mgt, 48.6% were held by male workers, and 51.4% were held by female workers. The other categories were: Svc, 50.6% male, 49.4% female; SO, 38.3% male, 61.7% female; NCM, 95.9% male, 4.1% female; PTM, 80.2% male, 19.8% female.

source: United States Department of Labor, Bureau of Labor Statistics report “Median weekly earnings of full-time wage and salary workers by detailed occupation and sex” covering the year 2010.
ftp://ftp.bls.gov/pub/special.requests/lf/aat39.txt

a) Put this data into a tree diagram.

b) What is the probability that a U.S. worker during 2010 was a man working in a sales or office job?

c) What percent of the U.S. workforce during 2010 was female?