Conditional Probability

 The following table was gathered from the National Diabetes Statistics Report, 2020

From the Center of Disease Control. We have some important definitions to consider in using this information.

# <https://www.cdc.gov/diabetes/library/features/diabetes-stat-report.html>

Conditional Probabilities are a fabulous tool in using probabilities to compare independent groups.

 **Probability of an event A given B**

$$P\left(A\left|B\right.\right)=\frac{P\left(A and B\right)}{P\left(B\right)}$$

 The following is a table created illustrating the results of the CDC report on the number of existing Diabetes cases with the number of undiagnosed cases for Diabetes. The table is illustrated below and everyone in the table has Diabetes (either type I or type 2).

# Age versus Diagnosis Type

|  | **Diagnosed**  | **Undiagnosed**  |  |
| --- | --- | --- | --- |
| **Age** | **Diabetes** | **Diabetes** | **Total** |
| **18 to 44** | 30 | 11 | 41 |
| **45 to 64** | 138 | 36 | 174 |
| **65 or older** | 214 | 54 | 268 |
| **Total**  | 382 | 101 | 483 |

If you select a person from this table at random, what’s the probability the person:

**Approximate to the Thousandths**

1. Diagnosed?
2. Diagnosed **given that** the person is aged 18 to 44 years?
3. Diagnosed **given that** the person is aged 45 to 64 years?
4. Diagnosed **given that** the person is 65 or older?
5. Which group was most likely to be diagnosed? Why?
6. Which group was least likely to be diagnosed? Why?
7. Undiagnosed?
8. Undiagnosed **given that** the person is aged 18 to 44 years?
9. Undiagnosed **given that** the person is aged 45 to 64 years?
10. Undiagnosed **given that** the person is 65 or older?
11. Which group was most likely to be undiagnosed? Why?
12. Which group was least likely to be undiagnosed? Why?

 The following is a table created illustrating the results of the CDC report on the number of existing Diabetes cases with the number of undiagnosed cases for Diabetes. The table is illustrated below and everyone in the table has Diabetes (either type I or type 2).

# Sex versus Diagnosis Type

|  | **Diagnosed**  | **Undiagnosed**  |  |
| --- | --- | --- | --- |
| **Sex** | **Diabetes** | **Diabetes** | **Total** |
| **Men** | 110 | 31 | 141 |
| **Women** | 95 | 25 | 120 |
| **Total**  | 205 | 56 | 261 |

Assuming this table represented a random sample, answer the following questions.

If you select a person at random, what’s the probability the person:

**Approximate to the Thousandths**

1. Diagnosed?
2. Diagnosed **given that** the person is a man?
3. Diagnosed **given that** the person is a woman?
4. Which group was most likely to be diagnosed? Why?
5. Which group was least likely to be diagnosed? Why?
6. Undiagnosed?
7. Undiagnosed **given that** the person is aged 18 to 44 years?
8. Undiagnosed **given that** the person is aged 45 to 64 years?
9. Undiagnosed **given that** the person is 65 or older?
10. Which group was most likely to be undiagnosed? Why?
11. Which group was least likely to be undiagnosed? Why?

 The following is a table created illustrating the results of the CDC report on the number of existing Diabetes cases with the number of undiagnosed cases for Diabetes. The table is illustrated below and everyone in the table has Diabetes (either type I or type 2).

# Race/Ethnicity versus Diagnosis Type

|  | **Diagnosed**  | **Undiagnosed**  |  |
| --- | --- | --- | --- |
| **Race-Ethnicity** | **Diabetes** | **Diabetes** | **Total** |
| **White (non-Hispanic)** | 94 | 25 | 119 |
| **Black (non-Hispanic)** | 133 | 30 | 163 |
| **Asian (Non-Hispanic)** | 112 | 46 | 158 |
| **Hispanic** | 103 | 35 | 138 |
| **Total**  | 442 | 136 | 578 |

Assuming this table represented a random sample, answer the following questions.

If you select a person at random, what’s the probability the person:

**Approximate to the Thousandths**

1. Diagnosed?
2. Diagnosed **given that** the person is White (non-Hispanic)?
3. Diagnosed **given that** the person is Black (non-Hispanic)?
4. Diagnosed **given that** the person is Asian (non-Hispanic)?
5. Diagnosed **given that** the person is Hispanic?
6. Which group was most likely to be diagnosed? Why?
7. Which group was least likely to be diagnosed? Why?
8. Undiagnosed?
9. Undiagnosed **given that** the person is White (non-Hispanic)?
10. Undiagnosed **given that** the person is Black (non-Hispanic)?
11. Undiagnosed **given that** the person is Asian (non-Hispanic)?
12. Undiagnosed **given that** the person is Hispanic?
13. Which group was most likely to be undiagnosed? Why?
14. Which group was least likely to be undiagnosed? Why?

 AB705 is a State Law that allows students to take transfer level English, and Mathematics courses, without first assessing the skills sets required to succeed in those courses. The new model assumes students have the necessary skill sets and claims that additional measures such as High Scholl GPA are better predictors of student success. AB705 mandated that all Community Colleges initiate the new assessment measures in Fall 2020. The following data was gathered from the Los Angeles Community College District on Ethnicity and student success in Math 227 for Fall 2020.

<https://assessment.cccco.edu/ab-705-implementation>

# Fall 2019 Math 227 Student Success Course Outcomes by Ethnicity

| **Ethnicity** | **Successful****A or B or C** | **D** | **F** | **W**  | **Drop** | **Total** |
| --- | --- | --- | --- | --- | --- | --- |
| **American Indian** | 8 | 1 | 2 | 2 | 0 | 13 |
| **Asian** | 388 | 18 | 53 | 97 | 1 | 557 |
| **Black** | 460 | 70 | 209 | 311 | 11 | 1061 |
| **Filipino** | 172 | 25 | 41 | 58 | 1 | 297 |
| **Hispanic** | 2652 | 505 | 1485 | 2124 | 41 | 6807 |
| **Multiethnic** | 127 | 8 | 37 | 81 | 3 | 256 |
| **Pacific Islander** | 11 | 2 | 6 | 9 | 0 | 28 |
| **White** | 532 | 50 | 123 | 229 | 11 | 945 |
| **Unknown** | 202 | 17 | 53 | 75 | 5 | 352 |
| **Total** | 4552 | 696 | 2009 | 2986 | 73 | 10316 |

If you select a person at random from this table, what’s the probability the person was: **Approximate to the Thousandths**

1. Successful?
2. Successful **given that** the person is Asian?
3. Successful **given that** the person is Black?
4. Successful **given that** the person is Hispanic?
5. Successful **given that** the person is White?
6. Failed?
7. Earned a D grade?
8. Unsuccessful (D or F or W or Drop)?
* Which ethnicity is more likely to be successful? Explain why?
* Which ethnicity is least likely to be successful? Explain why?

# Fall 2019 Math 227 Student Success Course Outcomes by Sex

| **Sex** | **Successful****A or B or C** | **D** | **F** | **W** | **Drop** | **Total** |
| --- | --- | --- | --- | --- | --- | --- |
| Female | 2879 | 431 | 1199 | 1767 | 46 | 6322 |
| Male | 1673 | 265 | 810 | 1219 | 27 | 3994 |
| **Total** | 4552 | 696 | 2009 | 2986 | 73 | 10316 |

If you select a person at random from this table, what’s the probability the person was: **Approximate to the Thousandths**

1. Successful?
2. Successful **given that** the student is a female?
3. Successful **given that** the person is a male?
4. Failed?
5. Unsuccessful (D or F or W or Dropped)?
6. Failed **given that** the person is a female?
7. Failed **given that** the person is a male?
8. Unsuccessful **given that** the person is a female?
9. Unsuccessful **given** that the person is a male?
* Which sex is more likely to be successful? Explain why?
* Which sex is more likely to be unsuccessful? Explain why?

# Fall 2018 Math 227 Student Success Course Outcomes by Ethnicity

| **Ethnicity** | **Successful** | **D** | **F** | **W** | **Drop** | **Total** |
| --- | --- | --- | --- | --- | --- | --- |
| **American Indian** | 3 | 0 | 4 | 3 | 0 | 10 |
| **Asian** | 314 | 14 | 42 | 65 | 2 | 437 |
| **Black** | 228 | 36 | 78 | 135 | 4 | 481 |
| **Filipino** | 120 | 9 | 23 | 38 | 1 | 191 |
| **Hispanic** | 1903 | 295 | 719 | 1029 | 10 | 3956 |
| **Multiethnic** | 61 | 7 | 19 | 26 | 0 | 113 |
| **Pacific Islander** | 6 | 0 | 1 | 5 | 0 | 12 |
| **White** | 363 | 28 | 45 | 123 | 3 | 562 |
| **Unknown** | 137 | 16 | 29 | 60 | 1 | 243 |
| **Total** | 3135 | 405 | 960 | 1484 | 21 | 6005 |

If you select a person at random from this table, what’s the probability the person was: **Approximate to the Thousandths**

1. Successful?
2. Successful **given that** the person is Asian?
3. Successful **given that** the person is Black?
4. Successful **given that** the person is Hispanic?
5. Successful **given that** the person is White?
6. Failed?
7. Earned a D grade?
8. Unsuccessful (D or F or W or Drop)?
* Which ethnicity is more likely to be successful? Explain why?
* Which ethnicity is least likely to be successful? Explain why?

# Fall 2018 Math 227 Student Success Course Outcomes by Sex

| **Sex** | **Successful****A or B or C** | **D** | **F** | **W** | **Drop** | **Total** |
| --- | --- | --- | --- | --- | --- | --- |
| Female | 1947 | 240 | 587 | 919 | 1760 | 3707 |
| Male | 1188 | 166 | 373 | 564 | 1110 | 2298 |
| Total | 3135 | 406 | 960 | 1483 | 2870 | 6005 |

If you select a person at random from this table, what’s the probability the person was: **Approximate to the Thousandths**

1. Successful?
2. Successful **given that** the student is a female?
3. Successful **given that** the person is a male?
4. Failed?
5. Unsuccessful (D or F or W or Dropped)?
6. Failed **given that** the person is a female?
7. Failed **given that** the person is a male?
8. Unsuccessful **given that** the person is a female?
9. Unsuccessful **given** that the person is a male?
* Which sex is more likely to be successful? Explain why?
* Which sex is more likely to be unsuccessful? Explain why?
* In which term were students more likely to fail? Explain why?
* In which term were students more likely to be successful? Explain why?

Quality of Education for STEM Careers

<https://www.pewsocialtrends.org/2018/01/09/women-and-men-in-stem-often-at-odds-over-workplace-equity/>

The following two tables compare the beliefs of Americans regarding the quality of Education of STEM Education at various levels in 2018.

# US Adults who rate the US when it comes to STEM Education at each level

|  | **Above Average** | **Average** | **Below Average** | **Total** |
| --- | --- | --- | --- | --- |
| **K12 Public Schools** | 25 | 44 | 31 | 100 |
| **Undergraduate** | 35 | 47 | 18 | 100 |
| **Graduate** | 38 | 44 | 18 | 100 |
| Total | 98 | 91 | 67 | 300 |

 72. Which level is more likely to be considered above average?

$$p(above average\left|K12), \right. p(above average\left|Undergraduate), \right. p(above average\left|Graduate) \right.$$

 73. Which level is more likely to be considered average?

 $p(average\left|K12), \right. p(average\left|Undergraduate), \right. p(average\left|Graduate) \right.$

 74. Which level is more likely to be considered below average?

 $p(below average\left|K12), \right. p(below average\left|Undergraduate), \right. p(below average)$

# US STEM Degree holders who rate the US when it comes to STEM Education at each Level

|  | **Above Average** | **Average** | **Below Average** | **Total** |
| --- | --- | --- | --- | --- |
| **K12 Public Schools** | 13 | 36 | 51 | 100 |
| **Undergraduate** | 52 | 35 | 13 | 100 |
| **Graduate** | 62 | 29 | 9 | 100 |
| Total | 127 | 100 | 73 | 300 |

 75. Which level is more likely to be considered above average?

 Compute $p(above average\left|K12), \right. p(above average\left|Undergraduate), \right. p(above average\left|Graduate) \right.$and compare likelihood.

 76. Which level is more likely to be considered average?

 Compute $p(average\left|K12), \right. p(average\left|Undergraduate), \right. p(average\left|Graduate) \right.$and compare likelihood.

 77. Which level is more likely to be considered below average?

 Compute $p(below average\left|K12), \right. p(below average\left|Undergraduate), \right. p(below average\left|Graduate) \right.$and compare likelihood.

Guns and Violence Against Women

<https://everytownresearch.org/reports/guns-intimate-partner-violence/>

 The following table represents Intimate Partner Violence (IPV) against women based on ethnicity. These are reported cases of violence by women of various ethnic groups committed by their intimate partners in 2019.

# Intimate Partner Violence (IPV)

|  | **Black Women/** | **American Indian** |  | **Other**  |  |
| --- | --- | --- | --- | --- | --- |
|  | **African American** | **Alaska Native Women** | **Hispanic** | **White Women** | **Total** |
| Experienced IPV | 55 | 4 | 49 | 239 | 19 | 366 |
| Not Experienced IPV | 69 | 4 | 83 | 444 | 34 | 634 |
| Total | 124 | 8 | 132 | 683 | 53 | 1000 |
|  |  |  |  |  |  |  |

If you select a woman at random, what’s the probability the woman:

78. Experienced IPV **given that** the woman is Black/African American?

79. Experienced IPV **given that** the woman is American Indian?

80. Experienced IPV **given that** the woman is Hispanic?

81. Experienced IPV **given that** the woman is White?

Which ethnic group is more likely to be victim of IPV?

False Positives and False Negatives; True Positives and True Negatives

 Unfortunately, people develop inflictions with various viruses and develop illnesses. We often seek to test individuals for those inflictions and illnesses. We also test individuals for various other medical conditions or for possible drug use. Every test has a certain degree of accuracy that we should be made aware of and consider. Tests have various error rates that are defined below to help us understand aspects of testing that we should be aware.

**Def False Positive (FPR)**

The likelihood of incorrectly testing + for an infliction when you actually are not infected.

$$P\left(test+\left|not Infected\right.\right)$$

**Def False Negative (FNR)**

The likelihood of incorrectly Testing – for an infliction when you are actually infected.

$$P\left(test-\left|nfected\right.\right)$$

**Def True Negative (TNR)**

The likelihood of correctly testing – when you really are not infected.

$$P\left(test-\left|not infected\right.\right)$$

**Def True Positive (TPR)**

The likelihood of correctly testing + when you are really infected.

$$P\left(test+\left|infected\right.\right)$$

**Def Prevalence**

The likelihood that a person has an illness or infliction.

$$P\left(infected\right)$$

False- Positive HIV Tests Results

The following table is based on information from the CDC for people who engage in “at risk” behavior and have a higher prevalence rate.

# <https://www.cdc.gov/hiv/pdf/testing/cdc-hiv-factsheet-false-positive-test-results.pdf>

|  | **Has HIV Virus** | **Not Have HIV Virus** | Total |
| --- | --- | --- | --- |
| Test + | 180 | 40 | 220 |
| Test - | 20 | 9760 | 9780 |
| Total | 200 | 9800 | 10000 |

Compute the following rates by selecting a person at random from this table.

82. False Positive Rate.

83. False Negative Rate.

84. True Positive Rate.

85. True Negative Rate.

Breast Cancer in Women

<https://brownmath.com/stat/falsepos.htm>

The following information was obtained from a study on the effectiveness of Core-Needle and Open Surgical biopsy for the diagnosis of Breast Lesions.

|  | **Have Breast Cancer** | **Doesn't Have Breast Cancer** | Total |
| --- | --- | --- | --- |
| **Test +** | 25800 | 1400 | 27200 |
| **Test -** | 4200 | 68600 | 72800 |
| Total | 30000 | 70000 | 100000 |
|  |  |  |  |

Compute the following rates by selecting a person at random from this table.

86. False Positive Rate.

87. False Negative Rate.

88. True Positive Rate.

89. True Negative Rate.

SARS-CoV-2 IgG antibody tests for COVID-19

The concepts of Sensitivity and Specificity for testing for Covid-19 vary by type of tests that are used. There are various types of tests that have particular benefits and drawbacks. You can read about some of them from information posted by ARUP Laboratories.

<https://www.aruplab.com/news/4-21-2020/How-Accurate-Are-COVID-19-Tests>

The following table summarized information found from ARUP Laboratories and represents a information about a test that detects IgG antibodies specific to the S1 domain of the spike protein of SARS-CoV-2.

<https://www.aruplab.com/infectious-disease/coronavirus/testing>

# Performance of IgG antibody test

|  | **Infected** | **not Infected** | **Total** |
| --- | --- | --- | --- |
| test + | 88 | 4 | 92 |
| test - | 0 | 1066 | 1066 |
| Total | 88 | 1070 | 1158 |

Compute the following rates by selecting a person at random from this table.

90. False Positive Rate.

91. False Negative Rate.

92. True Positive Rate.

93. True Negative Rate.

The following table summarized information found from ARUP Laboratories and represents a information about a test that detects IgG antibodies specific to the nucleocapsid protein of SARS-CoV-2.

# <https://www.aruplab.com/infectious-disease/coronavirus/testing>

|  | **Infected** | **not Infected** | **Total** |
| --- | --- | --- | --- |
| test + | 27 | 0 | 27 |
| test - | 3 | 80 | 83 |
| Total | 30 | 80 | 110 |

Compute the following rates by selecting a person at random from this table.

94. False Positive Rate.

95. False Negative Rate.

96. True Positive Rate.

97. True Negative Rate.