Probability

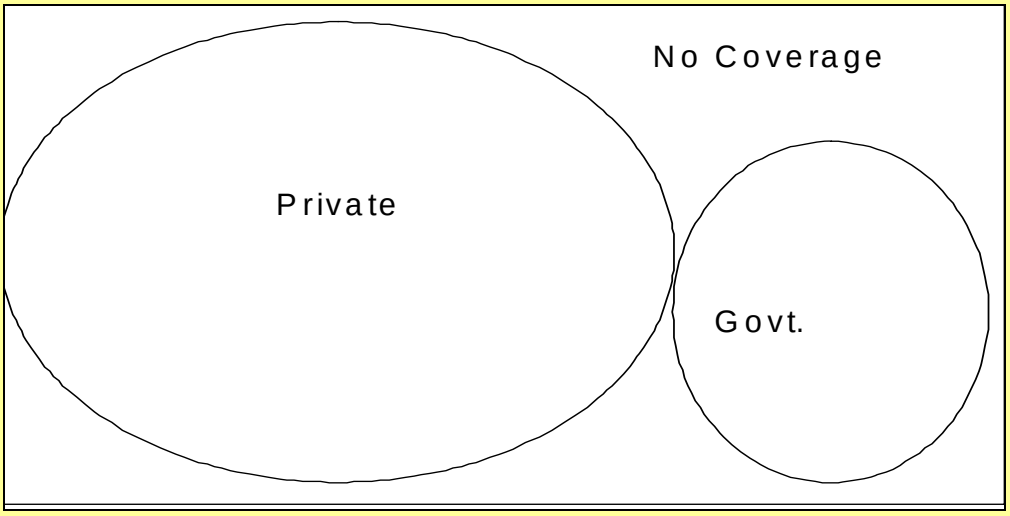
Probabilities are like relative frequencies for large samples--

Randomly pick 1 person—how likely are they to be in each group?
What proportion is in each group?

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?

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**	**1.00**



Probabilities for one "outcome set" are:

- Mutually exclusive (can be in only 1 category)
- Exhaustive (have to be in 1 category)
- Whole group must sum to 1

Use these to check your work

Total Probability for event:

You can always add mutually exclusive categories (cells in table).
Add up probabilities of different ways of having an "event".

What is the probability that a person in the sample had medical insurance?

P(Ins) = P(Pri)+P(Gov)=.85+.10=**.95**

What is the probability that a person had at least 1 doctor visit?

P(DV) = P(DV and Priv) + P(DV and Gov)+P(DV and No Cov) =

.72 + .09 + .03 = **.84**

Complement ("not" the event)

Can always find the complement since probabilities sum to 1.

What is the probability that a person had no medical insurance?

P(no Insurance)= 1 – P(Insurance)= 1 - .95 = **.05**

What is the probability that a person did not see a doctor during this two-year period?

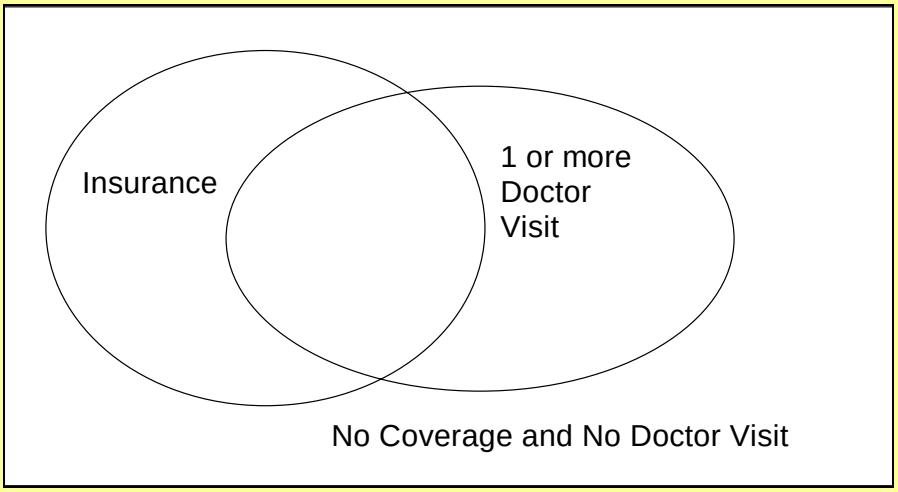
P(no DV)=1-P(DV)= 1 - .84 = **.16**

Intersections ("and" events) P(A and B)

What proportion had insurance and visited the doctor?

P(DV and Ins)= P(DV and Priv)+P(DV and Gov) =

= .72 +. 09 = **.81**



Union Rule ("OR" events)

Add probabilities of events and subtract double-counting

P(A or B) = P(A) + P(B) – P(A and B)

What is the probability of either visiting the doctor or having insurance?

P(DV or Ins) = P (DV) + P(Ins) – P(DV and Ins)

= .84+ .95 - .81 = **.98**

Conditional Probability: resets population (denominator of probability)
"among a subset of cases, what is the probability"

P(A|B) = P(A and B)/P(B)

Given that a person had insurance, what is the probability that the person visited the doctor?

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

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

P(DV | No Ins) = P(DV and no Ins)/P(no Ins)= .03/.05 = **.60**

Statistical Independence—events are not related to each other
Are conditional probabilities the same for each sub-group?

Extra Tricks:

Can reverse the conditional probability formula to find the intersection:

P(A and B) = P(A|B)P(B) = P(B|A)P(A)

(the multiplicative rule of probability)

What is the probability of someone having insurance and no visits?

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

=[1 – P(DV | Ins)]P(Ins)

=[1- .85](.95) = **.14**

Bayes Theorem—finding new conditional probabilities.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|not\,B)P(not\,B)}$$

What proportion of people with at least one doctor's visit did not have insurance?

P(no Ins | DV) =

= P(DV| no Ins)P(no Ins) /

[P(DV | no Ins)P(no Ins)+ P(DV | Ins)P(Ins)]

= (.60)(.05) / [(.60)(.05) + (.85)(.95)]

= .04

Table 1. Use of Health Care Services by Women: US, MEPS Data 2000

	Total Population	Any use of health care services (%)	Any ambulatory care visits (%)	Any inpatient stays (%)	Home health visits (%)	Prescription drug expenses (%)	Percent with preventive health care services
Total	107,310	91.0	81.7	11.0	2.8	75.9	93.5
Age (y)							
18–44	55,507	87.9	77.0	10.3	0.8	67.8	93.4
45–64	31,993	92.4	83.6	7.6	1.6	80.2	93.8
65–74	10,027	96.8	92.0	14.6	5.6	91.0	95.4
75–84	7,171	98.1	93.7	21.3	12.4	92.2	93.6
≥85	2,611	95.9	86.1	24.9	23.6	91.7	83.1
Marital status							
Married	57,068	92.1	83.1	11.3	1.6	76.2	95.1
Divorced/separated	14,743	90.5	83.3	9.7	2.8	78.6	92.8
Widowed	11,986	95.6	89.6	18.3	11.4	90.0	92.2
Never married	23,513	86.0	73.2	7.1	1.3	66.2	90.7
Race/ethnicity							
White/other (non-Hispanic)	83,133	93.3	84.3	10.9	2.9	79.4	94.6
Black (non-Hispanic)	13,025	84.4	73.0	12.9	3.1	66.1	89.4
Hispanic	11,152	81.3	72.0	9.1	2.1	61.2	90.2
Education (y)							
< 12	22,277	88.1	80.1	15.7	6.2	75.0	89.6
12	35,533	89.6	79.7	11.0	2.3	75.2	92.2
> 12	48,802	93.3	84.0	8.8	1.5	77.1	96.5
Census region							
Northeast	20,920	91.2	82.8	10.2	2.8	76.1	93.6
Midwest	24,378	93.8	85.9	12.6	3.3	79.2	94.9
South	38,297	89.6	80.5	11.8	2.7	76.1	92.7
West	23,715	90.0	78.3	8.7	2.5	71.9	93.1
Rural–Urban							
Metro	87,158	91.0	81.3	10.6	2.8	74.9	93.3
Near metro	7,796	91.3	82.8	11.7	2.5	80.0	92.7
Near rural	7,479	93.8	86.3	12.6	2.8	83.4	96.3
Rural	4,876	89.8	79.7	14.7	3.4	76.2	93.8
Health insurance status							
< 65 y							
Any private	68,123	92.1	81.9	7.9	0.6	75.1	96.2
Public only	8,018	93.5	86.0	26.4	5.6	80.0	90.4
Uninsured	11,360	71.2	59.7	5.7	0.9	50.8	80.2
> 65 y							
Medicare only	6,736	96.7	91.1	16.7	9.1	90.0	92.0
Medicare and private	10,579	97.6	93.1	18.7	8.4	93.0	94.7
Medicare and other public	2,396	96.6	90.3	21.8	23.8	91.9	90.0
Income							
Poor or near poor	17,146	87.7	79.0	17.9	5.9	74.1	87.8
Low income	14,505	88.7	81.1	15.3	5.0	74.5	90.0
Middle income	33,980	90.1	80.2	10.2	2.4	75.4	93.2
High income	41,679	93.8	84.2	7.3	1.2	77.5	97.3
Perceived health status							
< 65 y							
Excellent, very good or good	77,609	88.9	77.9	8.0	0.5	70.1	93.9
Fair or poor	9,816	94.6	91.5	19.2	5.6	90.5	91.7
≥ 65							
Excellent very good or good	14,400	96.7	91.3	13.1	6.7	90.3	95.9
Fair or poor	5,177	99.0	95.7	31.6	20.7	96.9	89.7

Note: Restricted to women 18 years and older in the civilian noninstitutionalized population. Percents may not add to one hundred due to rounding. Population estimates by education and health status exclude < 1.2% of cases due to item nonresponse.
Source: Center for Financing Access and Cost Trends, Agency for Healthcare and Quality; Medical Expenditure Panel Survey, 2000.