Using general addition rule compute the probability of a diamond or a face card?

Let $A$ be the event that a randomly selected card is a ♦.
Let $B$ be the event that a randomly selected card is a face card.

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

Note: $J$, $Q$, and $K$ ♦ fall into both categories
Using general addition rule compute the probability of a diamond or a face card?

Let $A$ be the event that a randomly selected card is a ♦.
Let $B$ be the event that a randomly selected card is a face card.

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

Note: J♦, Q ♦ and K ♦ fall into both categories
Note: Have to correct for double counting

\[ P(A) + P(B) = P(\spadesuit) + P(\text{face card}) = \frac{12}{52} + \frac{13}{52} \]

\[ P(A \cup B) = P(\text{face card or } \spadesuit) = P(\spadesuit) + P(\text{face card}) - P(\text{face card and } \spadesuit) \]

\[ = \frac{12}{52} + \frac{13}{52} - \frac{3}{52} = \frac{11}{26} \]
A *probability distribution* lists all possible events and the probabilities with which they occur.

* The probability distribution for the gender of one kid:

<table>
<thead>
<tr>
<th>Event</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>
**Probability distributions**

A *probability distribution* lists all possible events and the probabilities with which they occur.

* The probability distribution for the gender of one kid:

<table>
<thead>
<tr>
<th>Event</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

* Rules for probability distributions:
  * The events listed must be disjoint
  * Each probability must be between 0 and 1
  * The probabilities must total 1
A probability distribution lists all possible events and the probabilities with which they occur.

* The probability distribution for the gender of one kid:

<table>
<thead>
<tr>
<th>Event</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

* Rules for probability distributions:
  * The events listed must be disjoint
  * Each probability must be between 0 and 1
  * The probabilities must total 1

* The probability distribution for the genders of two kids:

<table>
<thead>
<tr>
<th>Event</th>
<th>MM</th>
<th>FF</th>
<th>MF</th>
<th>FM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>
In a survey, 52% of respondents said they are Democrats. What is the probability that a randomly selected respondent from this sample is a Republican?

* 0.48
* more than 0.48
* less than 0.48
* cannot calculate using only the information given
**Practice**

In a survey, 52% of respondents said they are Democrats. What is the probability that a randomly selected respondent from this sample is a Republican?

* 0.48
* more than 0.48
* less than 0.48
* cannot calculate using only the information given

If the only two political parties are Republican and Democrat, then (a) is possible. However it is also possible that some people do not affiliate with a political party or affiliate with a party other than these two. Then (c) is also possible. However (b) is definitely not possible since it would result in the total probability for the sample space being above 1.
**Sample space and complements**

*Sample space* is the collection of all possible outcomes of a trial.

* A couple has one kid, what is the sample space for the gender of this kid? 
  \[ S = \{M, F\} \]

* A couple has two kids, what is the sample space for the gender of these kids?
**Sample space and complements**

*Sample space* is the collection of all possible outcomes of a trial.

* A couple has one kid, what is the sample space for the gender of this kid? $S = \{M, F\}$
* A couple has two kids, what is the sample space for the gender of these kids? $S = \{MM, FF, FM, MF\}$
Sample space and complements

* A couple has one kid, what is the sample space for the gender of this kid? 
  \( S = \{M, F\} \)

* A couple has two kids, what is the sample space for the gender of these kids? 
  \( S = \{MM, FF, FM, MF\} \)

Complementary events are two mutually exclusive events whose probabilities that add up to 1.

* A couple has one kid. If we know that the kid is not a boy, what is gender of this kid? 
  \( \{F\} \rightarrow \text{Boy and girl are complementary outcomes.} \)

* A couple has two kids, if we know that they are not both girls, what are the possible gender combinations for these kids?
Sample space and complements

* Sample space is the collection of all possible outcomes of a trial.

  * A couple has one kid, what is the sample space for the gender of this kid? $S = \{M, F\}$

  * A couple has two kids, what is the sample space for the gender of these kids? $S = \{MM, FF, FM, MF\}$

Complementary events are two mutually exclusive events whose probabilities that add up to 1.

  * A couple has one kid. If we know that the kid is not a boy, what is gender of this kid? $\{♀, F\} \rightarrow$ Boy and girl are complementary outcomes.

  * A couple has two kids, if we know that they are not both girls, what are the possible gender combinations for these kids? $\{MM, FF, FM, MF\}$
Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.
INDEPENDENCE

Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.

* Knowing that the coin landed on a head on the first toss *does not* provide any useful information for determining what the coin will land on in the second toss. → Outcomes of two tosses of a coin are independent.
INDEPENDENCE

Two processes are *independent* if knowing the outcome of one provides no useful information about the outcome of the other.

* Knowing that the coin landed on a head on the first toss *does not* provide any useful information for determining what the coin will land on in the second toss. → Outcomes of two tosses of a coin are independent.

* Knowing that the first card drawn from a deck is an ace *does* provide useful information for determining the probability of drawing an ace in the second draw. → Outcomes of two draws from a deck of cards (without replacement) are dependent.
Between January 9-12, 2013, SurveyUSA interviewed a random sample of 500 NC residents asking them whether they think widespread gun ownership protects law abiding citizens from crime, or makes society more dangerous. 58% of all respondents said it protects citizens. 67% of White respondents, 28% of Black respondents, and 64% of Hispanic respondents shared this view. Which of the below is true?

Opinion on gun ownership and race ethnicity are most likely

* complementary
* mutually exclusive
* independent
* dependent
* disjoint

Between January 9-12, 2013, SurveyUSA interviewed a random sample of 500 NC residents asking them whether they think widespread gun ownership protects law abiding citizens from crime, or makes society more dangerous. 58% of all respondents said it protects citizens. 67% of White respondents, 28% of Black respondents, and 64% of Hispanic respondents shared this view. Which of the below is true?

Opinion on gun ownership and race ethnicity are most likely

* complementary
* mutually exclusive
* independent
* dependent
* disjoint

Checking for independence

If \( P(A \text{ occurs, given that } B \text{ is true}) = P(A \mid B) = P(A) \), then \( A \) and \( B \) are independent.
Checking for independence

If \( P(\text{A occurs, given that B is true}) = P(A \mid B) = P(A) \), then A and B are independent.

\[ P(\text{protects citizens}) = 0.58 \]
Checking for independence

If $P(A$ occurs, given that $B$ is true) = $P(A \mid B) = P(A)$, then $A$ and $B$ are independent.

$P($protects citizens$) = 0.58$

$P($randomly selected NC resident says gun ownership protects citizens, given that the resident is white$) = P($protects citizens $\mid$ White$) = 0.67$

$P($protects citizens $\mid$ Black$) = 0.28$

$P($protects citizens $\mid$ Hispanic$) = 0.64$
Checking for independence

If \( P(\text{A occurs, given that B is true}) = P(A \mid B) = P(A) \), then A and B are independent.

\[
P(\text{protects citizens}) = 0.58
\]

\[
P(\text{randomly selected NC resident says gun ownership protects citizens, given that the resident is white}) =
\]
\[
P(\text{protects citizens} \mid \text{White}) = 0.67
\]

\[
P(\text{protects citizens} \mid \text{Black}) = 0.28
\]

\[
P(\text{protects citizens} \mid \text{Hispanic}) = 0.64
\]

\[
P(\text{protects citizens}) \text{ varies by race/ethnicity, therefore opinion on gun ownership and race ethnicity are most likely dependent.}
\]
**Determining Dependence Based on Sample Data**

* If conditional probabilities calculated based on sample data suggest dependence between two variables, the next step is to conduct a hypothesis test to determine if the observed difference between the probabilities is likely or unlikely to have happened by chance.

* If the observed difference between the conditional probabilities is large, then there is stronger evidence that the difference is real.

* If a sample is large, then even a small difference can provide strong evidence of a real difference.
**Determining Dependence Based on Sample Data**

* If conditional probabilities calculated based on sample data suggest dependence between two variables, the next step is to conduct a hypothesis test to determine if the observed difference between the probabilities is likely or unlikely to have happened by chance.
* If the observed difference between the conditional probabilities is large, then there is stronger evidence that the difference is real.
* If a sample is large, then even a small difference can provide strong evidence of a real difference.

We saw that \( P(\text{protects citizens} \mid \text{White}) = 0.67 \) and \( P(\text{protects citizens} \mid \text{Hispanic}) = 0.64 \). Under which condition would you be more convinced of a real difference between the proportions of Whites and Hispanics who think gun widespread gun ownership protects citizens? \( n = 500 \) or \( n = 50,000 \).
Determining Dependence Based on Sample Data

* If conditional probabilities calculated based on sample data suggest dependence between two variables, the next step is to conduct a hypothesis test to determine if the observed difference between the probabilities is likely or unlikely to have happened by chance.

* If the observed difference between the conditional probabilities is large, then there is stronger evidence that the difference is real.

* If a sample is large, then even a small difference can provide strong evidence of a real difference.

We saw that \( P(\text{protects citizens} \mid \text{White}) = 0.67 \) and \( P(\text{protects citizens} \mid \text{Hispanic}) = 0.64 \). Under which condition would you be more convinced of a real difference between the proportions of Whites and Hispanics who think gun widespread gun ownership protects citizens? \( n = 500 \) or \( n = 50,000 \)

\( n = 50,000 \)
Product rule for independent events

\[ P(A \text{ and } B) = P(A) \times P(B) \]

Or more generally, \( P(A_1 \text{ and } \cdots \text{ and } A_r) = P(A_1) \times \cdots \times P(A_r) \)
Product rule for independent events

\[ P(A \text{ and } B) = P(A) \times P(B) \]

Or more generally, \( P(A_1 \text{ and } \cdots \text{ and } A_k) = P(A_1) \times \cdots \times P(A_k) \)

You toss a coin twice, what is the probability of getting two tails in a row?
Product rule for independent events

\[ P(A \text{ and } B) = P(A) \times P(B) \]

Or more generally, \( P(A_1 \text{ and } \cdots \text{ and } A_k) = P(A_1) \times \cdots \times P(A_k) \)

You toss a coin twice, what is the probability of getting two tails in a row?

\[ P(\text{T on the first toss}) \times P(\text{T on the second toss}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \]
A recent Gallup poll suggests that 25.5% of Texans do not have health insurance as of June 2012. Assuming that the uninsured rate stayed constant, what is the probability that two randomly selected Texans are both uninsured?

* $25.5^2$
* $0.255^2$
* $0.255 \times 2$
* $(1 - 0.255)^2$

A recent Gallup poll suggests that 25.5% of Texans do not have health insurance as of June 2012. Assuming that the uninsured rate stayed constant, what is the probability that two randomly selected Texans are both uninsured?

* $25.5^2$
* $0.255^2$
* $0.255 \times 2$
* $(1 - 0.255)^2$

Do the sum of probabilities of two disjoint events always add up to 1?
**Disjoint vs. Complementary**

Do the sum of probabilities of two disjoint events always add up to 1?

*Not necessarily, there may be more than 2 events in the sample space, e.g. party affiliation.*
Do the sum of probabilities of two disjoint events always add up to 1?

*Not necessarily, there may be more than 2 events in the sample space, e.g. party affiliation.*

Do the sum of probabilities of two complementary events always add up to 1?
**Disjoint vs. Complementary**

Do the sum of probabilities of two disjoint events always add up to 1?

*Not necessarily, there may be more than 2 events in the sample space, e.g. party affiliation.*

Do the sum of probabilities of two complementary events always add up to 1?

*Yes, that’s the definition of complementary, e.g. heads and tails.*
If we were to randomly select 5 Texans, what is the probability that at least one is uninsured?

* If we were to randomly select 5 Texans, the sample space for the number of Texans who are uninsured would be:

\[ S = \{0, 1, 2, 3, 4, 5\} \]

* We are interested in instances where at least one person is uninsured:

\[ S = \{0, 1, 2, 3, 4, 5\} \]

* So we can divide up the sample space into two categories:

\[ S = \{0, \text{at least one}\} \]
Putting everything together...

Since the probability of the sample space must add up to 1:

\[
\text{Prob(at least 1 uninsured)} = 1 - \text{Prob(none uninsured)}
\]
Since the probability of the sample space must add up to 1:

\[
Prob(\text{at least 1 uninsured}) = 1 - Prob(\text{none uninsured})
\]

\[
= 1 - [(1 - 0.255)^5]
\]
Since the probability of the sample space must add up to 1:

\[
\text{Prob(at least 1 uninsured)} = 1 - \text{Prob(none uninsured)}
\]
\[
= 1 - [(1 - 0.255)^5]
\]
\[
= 1 - 0.745^5
\]
Putting everything together...

Since the probability of the sample space must add up to 1:

\[
\text{Prob(\textit{at least 1 uninsured})} = 1 - \text{Prob(\textit{none uninsured})}
\]
\[
= 1 - [(1 - 0.255)^5]
\]
\[
= 1 - 0.745^5
\]
\[
= 1 - 0.23
\]
Since the probability of the sample space must add up to 1:

\[
\text{Prob(at least 1 uninsured)} = 1 - \text{Probnone uninsured)}
\]

\[
= 1 - [(1 - 0.255)^5]
\]

\[
= 1 - 0.745^5
\]

\[
= 1 - 0.23
\]

\[
= 0.77
\]

\[
P(\text{at least one}) = 1 - P(\text{none})
\]