Review for Exam 1 KEY

0.1 Probability Review

1. Jim and Mary are married.
   The probability that Jim watches TV after work is .6.
   The probability that Mary watches TV is .5.
   The probability that Jim watches TV given that Mary does is .8.

   (a) What is the probability that both Jim *and* Mary watch TV after work tonight?
   You have to use the general multiplication rule using the only conditional probability given to you, $P(\textbf{Jim}|\textbf{Mary})$.

   $$P(\textbf{Jim and Mary}) = P(\textbf{Jim}|\textbf{Mary}) \cdot P(\textbf{Mary}) = .8(.5) = .4$$

   (b) What is the probability that Mary watches TV tonight if Jim turns it on and sits down to watch?
   You need to use the conditional probability formula.

   $$P(\textbf{Mary}|\textbf{Jim}) = \frac{P(\textbf{Mary and Jim})}{P(\textbf{Jim})} = \frac{.4}{.6} = .6666$$

   (c) What is the probability that either Jim *or* Mary watch TV tonight?

   $$P(\textbf{Mary or Jim}) = P(Mary) + P(Jim) - P(\textbf{Mary and Jim}) = .5 + .6 - .4 = .7$$

2. You decide to roll 5 six sided fair dice. What is the probability that you get at least one 3 out of the five dice?
   The probability of getting a three on any dice is $1/6$. This means that the probability of getting something besides a three on any dice is $5/6$.
   It is much easier to use the complement rule for this problem. The complement of "at least one 3" is "no threes".

   $$\begin{aligned} P(\textbf{at least one three}) &= 1 - P(\textbf{no threes}) \\ &= 1 - P(\textbf{1st dice not a 3}) \cdot P(\textbf{2nd dice not a 3}) \cdots P(\textbf{5th dice not a 3}) \\ &= 1 - \left(\frac{5}{6}\right) \cdot \left(\frac{5}{6}\right) \cdot \left(\frac{5}{6}\right) \cdot \left(\frac{5}{6}\right) \cdot \left(\frac{5}{6}\right) \\ &= 1 - \left(\frac{5}{6}\right)^5 \\ &= .5981 \end{aligned}$$

3. A contingency table for injuries in the United States is shown below. Frequencies are in millions.

	Work	Home	Other	**Total**
Male	8	9.8	17.8	**35.6**
Female	1.3	11.6	12.9	**25.8**
Total	**9.3**	**21.4**	**30.7**	**61.4**

(a) What is the probability an injury occurs at home?
There are 21.4 injuries at home out of a total 61.4 injuries.

$$P\left(home\right) = \frac{21.4}{61.4} = .3485$$

(b) Find $P\left(\text{Female}\right)$.
There are 25.8 female injuries out of a total 61.4 injuries.

$$P\left(female\right) = \frac{25.8}{61.4} = .4202$$

(c) What is the probability that an injury occurs at home, if you know the person was a female?

$$P\left(\textbf{home}|\textbf{female}\right) = \frac{P\left(\textbf{home and female}\right)}{P\left(\textbf{female}\right)} = \frac{11.6/61.4}{25.8/61.4} = \frac{11.6}{25.8} = .4496$$

(d) Are the events "injury at home" and "injury to a female" independent? Justify your answer by comparing two probabilities.
You have only found one conditional probability so far involving home and female, $P\left(\textbf{home}|\textbf{female}\right)$, so you have to compare $P\left(\textbf{home}|\textbf{female}\right)$ and $P\left(\textbf{home}\right)$.

$$P\left(\textbf{home}|\textbf{female}\right) = \frac{P\left(\textbf{home and female}\right)}{P\left(\textbf{female}\right)} = \frac{11.6/61.4}{25.8/61.4} = \frac{11.6}{25.8} = .4496$$

$$P\left(\textbf{home}\right) = \frac{21.4}{61.4} = .3485$$

Since the probability that a person gets injured at home increases if you know the person is female, the events are DEPENDENT.

(e) Are you more likely to be injured at work if you are a male or female? What two probabilities should you compare?
You need to compare $P\left(Work \mid Male\right)$ and $P\left(Work \mid Female\right)$.

$$P\left(Work \mid Male\right) = \frac{P\left(\textbf{Work and Male}\right)}{P\left(Male\right)} = \frac{8.0/61.4}{35.6/61.4} = \frac{8.0}{35.6} = .2247$$

$$P\left(Work \mid Female\right) = \frac{P\left(\textbf{Work and Female}\right)}{P\left(Female\right)} = \frac{1.3/61.4}{25.8/61.4} = \frac{1.3}{25.8} = .0504$$

If you are male, the probability of being injured at work is 22%; if you are female, the probability of being injured at work is 5%. So you are more likely to be injured at work if you are male. Why do you think that is?

(f) Find $P\left(\text{Male and Work}\right)$.
There are 8 injuries that happen at work and to males out of the 61.4 injuries.

$$P\left(\textbf{Male and Work}\right) = \frac{8}{61.4} = .1303$$

(g) Are the events "Male" and "Work" disjoint? Give a numerical justification.
There are 8 million injuries that happened at work to males. So obviously it is possible for an injury to happen to a male and at work at the same time. So they are not disjoint!
Also, $P\,(\textbf{Male and Work}) = .1303$ **instead of zero. The intersection of disjoint events has a probability of zero.**

(h) Find $P\,(\text{Male or Work})$.
You need to make sure that you don't double count the intersection.

$$\begin{aligned} P\,(\textbf{Male or Work}) &= P\,(Male) + P\,(Work) - P\,(Male \textbf{ and } Work) \\ &= \frac{35.6}{61.4} + \frac{9.3}{61.4} - \frac{8}{61.4} \\ &= .6010 \end{aligned}$$

(i) Find $P\,(\text{Female}|\text{Home})$.
This means the probability of the injury occurring to a female, if we know it happened at home.

$$P\,(\textbf{Female}|\textbf{Home}) = \frac{P\,(\textbf{Female and Home})}{P\,(\textbf{Home})} = \frac{11.6/61.4}{21.4/61.4} = \frac{11.6}{21.4} = .5421$$

(j) Based on your gender, where are you more likely to be injured? (Hint: Do this for your specific gender. You will need to compute 3 probabilities.)
If you are Male, you are most likely to be injured at Other:

$$P\,(\textbf{Work}|\textbf{Male}) = \frac{8.0}{35.6} = .2247$$

$$P\,(\textbf{Home}|\textbf{Male}) = \frac{9.8}{35.6} = .2753$$

$$P\,(\textbf{Other}|\textbf{Male}) = \frac{17.8}{35.6} = .5000$$

If you are Female, you are most likely to be injured at Other:

$$P\,(\textbf{Work}|\textbf{Female}) = \frac{1.3}{25.8} = .0504$$

$$P\,(\textbf{Home}|\textbf{Female}) = \frac{11.6}{25.8} = .4496$$

$$P\,(\textbf{Other}|\textbf{Female}) = \frac{12.9}{25.8} = .5000$$

4. For space shuttles, NASA standards state that each critical component of the mission must have 99.99% reliability, or that the probability of failure is only .0001. The space mission on which the Challenger exploded had 748 critical components.

   (a) What is the probability that none of the critical components would fail?
   The probability of a component not failing is .9999. We need all 748 components to not fail.

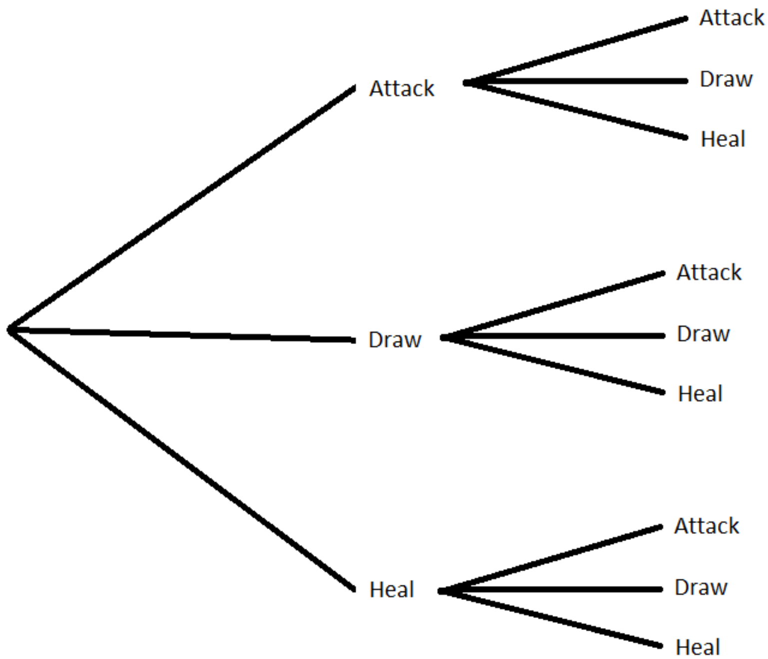
$$\begin{aligned} P\left(\textbf{no components fail}\right) &= (.9999)\cdot(.9999)\cdot(.9999)\cdot(.9999)\cdot\cdots(.9999) \\ &= (.9999)^{748} \\ &= .9279 \end{aligned}$$

   (b) What is the probability that at least one critical component would fail?
   You need the complement rule. The complement of at least one failure is no failures.

$$\begin{aligned} P\left(\textbf{at least one failure}\right) &= 1 - P\left(\textbf{no components fail}\right) \\ &= 1-(.9999)\cdot(.9999)\cdot(.9999)\cdot(.9999)\cdot\cdots(.9999) \\ &= 1-(.9999)^{748} \\ &= 1-.9279 \\ &= .0721 \end{aligned}$$

5. In the game Epic Duels, a person gets two actions per turn. For each action they can attack, draw a card, or heal. Suppose attacking, drawing a card, and healing are equally likely.

   (a) What is the sample space?
   The sample space is S={AA,AD,AH,DA,DD,DH,HA,HD,HH}.



   (b) What is the probability of each possible outcome?
   Because there are 9 possible outcomes that are equally likely, the probability of each one is 1/9.

6. Is it more likely to get 4 heads out of 5 coin tosses or 400 heads out of 500 coin tosses?
   There is always more variability for fewer trials. So it is very plausible to get 4 heads out of 5 coin tosses, but by the time we get to 500 coin tosses, the percentage of heads should be very close to 50%.

7. Here is the information for a class on campus. A student can't be a freshman and a sophomore at the same time, etc.

Class	Frequency
Freshman	21
Sophomore	14
Junior	9
Senior	6

(a) What is the probability that a randomly selected student is a Sophomore?
There are 14 Sophomores out of 50 students.

$$P\left(Sophomore\right) = \frac{14}{50} = .28$$

(b) Are the events "Sophomore" and "Junior" disjoint?
Yes, because you can't be a Sophomore and a Junior at the same time.
This also means that $P\left(\textbf{Sophomore and Junior}\right) = 0$.

(c) What is the probability that a randomly selected student is a Sophomore or a Junior?

$$P\left(\textbf{Sophomore or Junior}\right) = P\left(Sophomore\right) + P\left(Junior\right) = \frac{14}{50} + \frac{9}{50} = \frac{23}{50} = .46$$

(d) What is the probability that a randomly selected student is not a Senior?
Method 1: (The long way)

$$\begin{aligned} P\left(\textbf{not Senior}\right) &= P\left(Freshman\right) + P\left(Sophomore\right) + P\left(Junior\right) \\ &= \frac{21}{50} + \frac{14}{50} + \frac{9}{50} \\ &= \frac{44}{50} \\ &= .88 \end{aligned}$$

Method 2: (The complement rule)

$$\begin{aligned} P\left(\textbf{not Senior}\right) &= 1 - P\left(Senior\right) \\ &= 1 - \frac{6}{50} \\ &= \frac{44}{50} \\ &= .88 \end{aligned}$$

(e) We hold a drawing for prizes at the end of the semester. We give out 3 prizes to randomly selected students and students can be chosen more than once.

i. Are the three student selections going to be independent?
Since the student can be chosen more than once, we are sampling with replacement. So the choice of the first student won't affect the choice of the second student. So they are independent.

ii. What is the probability that we choose a Freshman, then a Sophomore, then a Junior?

$$\begin{aligned} P\left(\textbf{Freshman and Sophomore and Junior}\right) &= P\left(Freshman\right) \cdot P\left(Sophomore\right) \cdot P\left(Junior\right) \\ &= \frac{21}{50} \cdot \frac{14}{50} \cdot \frac{9}{50} \\ &= .0212 \end{aligned}$$

iii. What is the probability that we choose three Seniors?

$$\begin{aligned} P\left(\textbf{Senior and Senior and Senior}\right) &= P\left(Senior\right)\cdot P\left(Senior\right)\cdot P\left(Senior\right) \\ &= \frac{6}{50}\cdot\frac{6}{50}\cdot\frac{6}{50} \\ &= .0017 \end{aligned}$$

iv. What is the probability that we choose at least one Senior?

It is easier to use the complement rule than to figure out all the possible ways we can get at least one Senior (e.g. Senior&Senior&Freshman or Senior&Senior&Junior ...)

The complement of at least one Senior is no seniors. So we want to get something besides a senior each time. For one draw there are 44 students that aren't seniors out of 50 total students.

$$\begin{aligned} P\left(\textbf{at least one Senior}\right) &= 1 - P\left(\textbf{no Seniors}\right) \\ &= 1 - P\left(\textbf{1st not senior}\right)\cdot\left(\textbf{2nd not senior}\right)\cdot\left(\textbf{3rd not senior}\right) \\ &= 1 - \frac{44}{50}\cdot\frac{44}{50}\cdot\frac{44}{50} \\ &= 1 - .6815 \\ &= .3185 \end{aligned}$$

8. Here is the information for a class on campus. We need to randomly select two students to interview. (We don't want to interview the same student twice.)

Class	Frequency
Freshman	21
Sophomore	14
Junior	9
Senior	6

(a) Are the two student selections going to be independent?
Since we don't want to interview the same student twice, we are sampling without replacement. So the first student we select will affect our options for the second student. So they are DEPENDENT.

(b) What is the probability that we choose a Freshman and then a Senior?
First we choose a freshman and there are 21 freshman out of 50 students. But once we pick a freshman then there are 6 seniors out of 49 students.

$$\begin{aligned} P\left(\textbf{freshman and then senior}\right) &= P\left(freshman\right) \cdot P\left(senior \mid \textbf{already chose freshman}\right) \\ &= \frac{21}{50} \cdot \frac{6}{49} \\ &= .0514 \end{aligned}$$

(c) What is the probability that we choose a Sophomore and then another Sophomore?

$$\begin{aligned} P\left(\textbf{Sophomore and then Sophomore}\right) &= P\left(Sophomore\right) \cdot P\left(Sophomore \mid \textbf{already chose a Sophomore}\right) \\ &= \frac{14}{50} \cdot \frac{13}{49} \\ &= .0743 \end{aligned}$$

(d) What is the probability that we choose a Freshman and a Senior in any order?
There are two ways we can get a Freshman and a Senior. We can have a (Freshman and then a Senior) or a (Senior and then a Freshman).

$$\begin{aligned} P\left(\texttt{Freshman and Senior}\right) &= P\left(\texttt{Freshman and then Senior}\right) + P\left(\texttt{Senior and then Freshman}\right) \\ &= \frac{21}{50} \cdot \frac{6}{49} + \frac{6}{50} \cdot \frac{21}{49} \\ &= .1029 \end{aligned}$$

(e) What is the probability that both students are at the same class level?
To achieve this, we can have a Freshman&Freshman or Sophomore&Sophomore or Junior&Junior or Senior&Senior.

$$\begin{aligned} P\left(\textbf{both same level}\right) &= P\left(\textbf{Fresh and Fresh}\right) + P\left(\textbf{Soph and Soph}\right) + P\left(\textbf{Jun and Jun}\right) + P\left(\textbf{Sen and Sen}\right) \\ &= \frac{21}{50} \cdot \frac{20}{49} + \frac{14}{50} \cdot \frac{13}{49} + \frac{9}{50} \cdot \frac{8}{49} + \frac{6}{50} \cdot \frac{5}{49} \\ &= \frac{704}{2450} \\ &= .2873 \end{aligned}$$

(f) What is the probability that both students are not at the same class level?
You could list out all the possibilities such as Freshman&Sophomore or Freshman&Junior, etc. But this would take a while. It is easier to use the complement rule. The complement of "not the same level" is "both same level".

$$\begin{aligned} P\left(\textbf{not the same level}\right) &= 1 - P\left(\textbf{both same level}\right) \\ &= 1 - .2873 \\ &= .7127 \end{aligned}$$

(g) We decide we need to interview more students. We decide to interview a total of 6 randomly chosen students. What is the probability that all the students chosen are freshmen?

$$P\left(\textbf{6 freshmen}\right) = \frac{21}{50} \cdot \frac{20}{49} \cdot \frac{19}{48} \cdot \frac{18}{47} \cdot \frac{17}{46} \cdot \frac{16}{45} = .0034$$

9. The National Safety council reported that 14.1% of drivers responsible for fatal crashes are between 16-20 years old; of those drivers, 12.7% had an elevated blood alcohol level.

   (a) Which of the following probability symbols represents the 12.7%?

      i. $P(16-20)$

      ii. $P(\text{blood alcohol})$

      iii. $P(16-20 \mid \text{blood alcohol})$

      iv. $P(\text{blood alcohol} \mid 16-20)$

      v. $P(\text{16-20 and blood alcohol})$

      We already know that the drivers are 16-20 years old and then 12.7% have elevated blood alcohol levels. So the answer is $P(\textbf{blood alcohol}|\textbf{16-20})$.

   (b) Find $P(\text{blood alcohol and 16-20})$.

   We need to use the general multiplication rule.

$$\begin{aligned} P(\textbf{16-20 and blood alcohol}) &= P(\textbf{blood alcohol}|\textbf{16-20}) \cdot P(16-20) \\ &= .127\,(.141) \\ &= .0179 \end{aligned}$$

   (c) We know that the probability that a driver responsible for a fatal crash is between 21 and 24 year old is .114. We also know that the probability that a driver responsible for a fatal crash is between 21-24 years old *and* has an elevated blood alcohol level is .0317.

   A crash occurred. We discover that the driver was between 21-24 year old. What is the probability that the driver had an elevated blood alcohol level?

$$\begin{aligned} P(\textbf{blood alcohol}|\textbf{21-24}) &= \frac{P(\textbf{blood alcohol and 21-24})}{P(21-24)} \\ &= \frac{.0317}{.114} \\ &= .2781 \end{aligned}$$

   (d) Is it more likely that a driver responsible for a fatal crash has an elevated blood alcohol level if they are between 16-20 or if they are between 21-24 years old? (What two probabilities do you need to compare?)

   You need to compare the probabilities $P(\textbf{blood alcohol}|\textbf{16-20})$ and $P(\textbf{blood alcohol}|\textbf{21-24})$.

$$P(\textbf{blood alcohol}|\textbf{16-20}) = .127$$

$$P(\textbf{blood alcohol}|\textbf{21-24}) = .2781$$

   We see that drivers that cause fatal crashes that are between 21-24 are more likely to have elevated blood alcohol levels than drivers aged 16-20. (This makes sense since the legal drinking age is 21.)

10. According to the Medical College of Wisconsin, 9% of men are color blind. If we randomly select four men, find the probability that:

(a) all of the men are color blind.

$$\begin{aligned} P\,(\textbf{all color blind}) &= P\,(CB)\cdot P\,(CB)\cdot P\,(CB)\cdot P\,(CB) \\ &= .09\,(.09)\,(.09)\,(.09) \\ &= .000066 \end{aligned}$$

(b) at least one man is color blind.
You need to use the complement rule. The complement of "at least one color blind man" is "no color blind men". The probability that one man is not color blind is $P\,(N) = 1 - .09 = .91$.

$$\begin{aligned} P\,(\textbf{at least one color blind}) &= 1 - P\,(\textbf{no color blind}) \\ &= 1 - P\,(N)\cdot P\,(N)\cdot P\,(N)\cdot P\,(N) \\ &= 1 - (.91)\,(.91)\,(.91)\,(.91) \\ &= 1 - .6857 \\ &= .3143 \end{aligned}$$

(c) the first three men are not color blind and the last man is color blind.

$$\begin{aligned} P(N,N,N,CB) &= (.91)\,(.91)\,(.91)\,(.09) \\ &= .0678 \end{aligned}$$

(d) exactly one of the men is color blind.
You have to figure out which sample space outcomes would correspond to exactly one color blind man. They are (CB,N,N,N), (N,CB,N,N), (N,N,CB,N), or (N,N,N,CB).

$$\begin{aligned} P\,(\textbf{one color blind}) &= P(CB,N,N,N) + P(N,CB,N,N) + P(N,N,CB,N) + P(N,N,N,CB) \\ &= (.09)\,(.91)\,(.91)\,(.91) + (.91)\,(.09)\,(.91)\,(.91) \\ &\quad + (.91)\,(.91)\,(.09)\,(.91) + (.91)\,(.91)\,(.91)\,(.09) \\ &= .0678 + .0678 + .0678 + .0678 \\ &= .2712 \end{aligned}$$

11. Cards numbered 1,2,3,...,10 are placed in a box. The box is shaken and a person selects two cards with replacement.

    (a) Are the card draws independent?
    Yes, because we put the card back before drawing the next card, so the first draw doesn't affect the second draw.

    (b) What is the probability that the first card is a 6?
    There are 10 cards and only one card with a 6.

    $$P\left(\textbf{card is a } 6\right) = \frac{1}{10}$$

    (c) If the first card is a 6, what is the probability of that the second card is a 9?
    Because we put the card back, we still have 10 cards to choose from. We don't need a conditional probability for independent events.

    $$P\left(\textbf{2nd card is a 9}\right) = \frac{1}{10}$$

    (d) What is the probability of the first card being a 6 and the second card being a 9?
    We need to use the general multiplication rule.

    $$\begin{aligned} P\left(\textbf{1st card is 6 AND 2nd card is 9}\right) &= P\left(\textbf{1st card is a 6}\right) \cdot P\left(\textbf{2nd card is 9}\right) \\ &= \frac{1}{10} \cdot \frac{1}{10} \\ &= \frac{1}{100} \end{aligned}$$

    (e) What is the probability that both cards selected have the numbers 6, 7, 8, 9, or 10?
    For the first draw, we have 5 possibilities out of 10 cards. For the second draw, we still have 5 possibilities out of 10 cards.

    $$\begin{aligned} P\left(\textbf{both cards have 6-10}\right) &= P\left(\textbf{1st card has 6-10}\right) \cdot P\left(\textbf{2nd card has 6-10}\right) \\ &= \frac{5}{10} \cdot \frac{5}{10} \\ &= \frac{1}{4} \end{aligned}$$

12. Cards numbered 1,2,3,...,10 are placed in a box. The box is shaken and a person selects two cards without replacement.

(a) Are the card draws independent?
No, because we didn't the card back before drawing the next card, the first draw affects the second draw.

(b) What is the probability that the first card is a 6?
There are 10 cards and only one card with a 6.

$$P\left(\textbf{card is a } 6\right) = \frac{1}{10}$$

(c) If the first card is a 6, what is the probability of that the second card is a 9?
Because we didn't put the card back, we only have 9 cards to choose from.

$$P\left(\textbf{2nd card is a 9}|\textbf{we already drew a 6}\right) = \frac{1}{9}$$

(d) What is the probability of the first card being a 6 and the second card being a 9?
We need to use the general multiplication rule.

$$\begin{aligned} P\left(\textbf{1st card is 6 AND 2nd card is 9}\right) &= P\left(\textbf{1st card is a 6}\right) \cdot P\left(\textbf{2nd card is 9}|\textbf{1st card is 6}\right) \\ &= \frac{1}{10} \cdot \frac{1}{9} \\ &= \frac{1}{90} \end{aligned}$$

(e) What is the probability that both cards selected have the numbers 6, 7, 8, 9, or 10?
For the first draw, we have 5 possibilities out of 10 cards. For the second draw, we have 4 possibilities out of 9 cards.

$$\begin{aligned} P\left(\textbf{both cards have 6-10}\right) &= P\left(\textbf{1st card has 6-10}\right) \cdot P\left(\textbf{2nd card has 6-10} \mid \textbf{1st card has 6-10}\right) \\ &= \frac{5}{10} \cdot \frac{4}{9} \\ &= \frac{2}{9} \\ &= .22222222222222 \end{aligned}$$

13. The Gallup Organization conducted a study of U.S. adults on their attitudes toward the death penalty. Here are the results.

	In Favor	Not in Favor	Not Sure	**Total**
Northeast	68	29	3	**100**
Midwest	69	29	2	**100**
South	72	23	5	**100**
West	80	19	1	**100**
Total	**289**	**100**	**11**	**400**

(a) Suppose we randomly select one U.S. adult. Find each probability.

i. $P\left(\text{In Favor}\right)$

There are 289 people in favor of the death penalty and 400 total people.

$$P\left(\textbf{In favor}\right) = \frac{289}{400} = .7225$$

ii. The person doesn't say "not sure".

We need to use the complement rule. There are 11 people that say "not sure".

$$P\left(\textbf{not sure}\right) = \frac{11}{400}$$

$$P\left(\textbf{don't say "not sure"}\right) = 1 - \frac{11}{400} = \frac{389}{400}$$

iii. The person is from the Northeast.

There are 100 people from the northeast.

$$P\left(\textbf{northeast}\right) = \frac{100}{400} = .25$$

iv. $P\left(\text{Northeast or In Favor}\right)$

$$\begin{aligned} P\left(\textbf{Northeast or In Favor}\right) &= P\left(\textbf{Northeast}\right) + P\left(\textbf{In Favor}\right) - P\left(\textbf{Northeast and In Favor}\right) \\ &= \frac{100}{400} + \frac{289}{400} - \frac{68}{400} \\ &= \frac{321}{400} \\ &= .8025 \end{aligned}$$

v. $P\left(\text{In Favor}|\text{Northeast}\right)$

You need to use the conditional probability formula.

$$\begin{aligned} P\left(\textbf{In favor}|\textbf{Northeast}\right) &= \frac{P\left(\textbf{In Favor and Northeast}\right)}{P\left(\textbf{Northeast}\right)} \\ &= \frac{68/400}{100/400} \\ &= \frac{68}{100} \\ &= .68 \end{aligned}$$

vi. $P\left(\text{Northeast}|\text{In Favor}\right)$

$$\begin{aligned} P\left(\textbf{Northeast}|\textbf{In Favor}\right) &= \frac{P\left(\textbf{In Favor and Northeast}\right)}{P\left(\textbf{In Favor}\right)} \\ &= \frac{68/400}{289/400} \\ &= \frac{68}{289} \\ &= .2352 \end{aligned}$$

vii. $P\,(\text{In Favor}|\text{South})$

$$\begin{aligned} P\,(\textbf{In Favor}|\textbf{South}) &= \frac{P\,(\textbf{In Favor and South})}{P\,(\textbf{South})} \\ &= \frac{72/400}{100/400} \\ &= \frac{72}{100} \\ &= .72 \end{aligned}$$

viii. $P\,(\text{South}|\text{In Favor})$

$$\begin{aligned} P\,(\textbf{South}|\textbf{In Favor}) &= \frac{P\,(\textbf{In Favor and South})}{P\,(\textbf{In Favor})} \\ &= \frac{72/400}{289/400} \\ &= \frac{72}{289} \\ &= .2491 \end{aligned}$$

ix. $P\,(\text{In Favor and West})$

There are 80 people who are in favor and live in the West.

$$P\,(\textbf{In Favor and West}) = \frac{80}{400} = .2$$

Of course you could use the general multiplication rule, but you don't have to since you can get the answer from the table.

(b) Is a person more likely to be in favor of the death penalty if they are in the Northeast or the South? Justify your answer.

You have to compare the two conditional probabilities:

$$P\,(\textbf{In Favor}|\textbf{South}) = .72$$

$$P\,(\textbf{In favor}|\textbf{Northeast}) = .68$$

If someone is from the south, there is a 72% chance that they are in favor of the death penalty.

If someone is from the northeast, there is a 68% chance that they are in favor of the death penalty.
So people from the south are more likely to be in favor of the death penalty.

(c) Are the events "Northeast" and "In Favor" independent? Justify your answer.

You can compare $P\,(\textbf{In favor}|\textbf{Northeast}) = .68$ **and** $P\,(\textbf{In favor}) = .7225$**. Since the probabilities aren't equal, we see that whether a person is from the Northeast affects whether they are in favor of the death penalty. Therefore they are DEPENDENT.**

You could also compare $P\,(\textbf{Northeast}|\textbf{In Favor}) = .2352$ **and** $P(Northeast) = .25$**. Since the probabilities are different, we know that the events are DEPENDENT.**

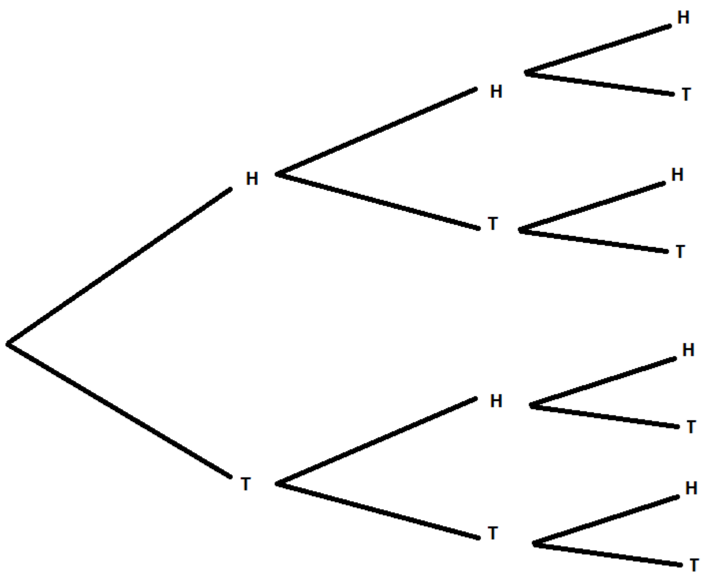
(d) Are the events "In Favor" and "West" disjoint? Justify your answer.

There are 80 people who are in favor and in the west. So obviously it is possible to be in favor and in the west at the same time. So they can't be disjoint.

Also, we know that if they were disjoint then $P\,(\textbf{In Favor and West})$ **would have to be ZERO. But** $P\,(\textbf{In Favor and West}) = \frac{80}{400} = .2$**. So the events are not disjoint.**

14. Suppose we toss a biased coin 3 times. This biased coin has 0.7 probability of being a head on any coin toss.

 (a) What is the sample space? (Hint: draw a tree diagram.)



S={HHH,HHT,HTH,HTT,THH,THT,TTH,TTT}

 (b) What is the probability of getting a tail on one coin toss?
 We have to use the complement rule.

$$\begin{aligned} P\,(tail) &= 1 - P\,(head) \\ &= 1 - 0.7 \\ &= 0.3 \end{aligned}$$

 (c) Should the results of the three coin tosses be independent?
 Yes!

 (d) Fill in the probability table for the possible outcomes:

Outcome	(Show your work)	Probability
HHH		
HHT		
HTH		
HTT		
THH		
THT		
TTH		
TTT		

You need to use the multiplication rule for independent events.

Outcome	(Show your work)	Probability
HHH	$P(HHH) = P(H) \cdot P(H) \cdot P(H) = (0.7)(0.7)(0.7) = 0.343$	0.343
HHT	$(.7)(.7)(.3)$	**0.147**
HTH	$(.7)(.3)(.7)$	**0.147**
HTT	**(.7)(.3)(.3)**	**0.063**
THH	**(.3)(.7)(.7)**	**0.147**
THT	**(.3)(.7)(.3)**	**0.063**
TTH	**(.3)(.3)(.7)**	**0.063**
TTT	**(.3)(.3)(.3)**	**0.027**

(e) Find the probabilities for the following events:

i. we get exactly 2 heads.
The outcomes that cover exactly two heads are HHT, HTH, or THH.

$$P(\textbf{exactly two heads}) = P(\textbf{HHT or HTH or THH}) = 0.147 + 0.147 + 0.147 = 0.441$$

ii. we get at least 2 heads.
The outcomes that cover at least two heads are HHT, HTH, and THH, or HHH.

$$P(\textbf{at least two heads}) = P(\textbf{HHT or HTH or THH or HHH}) = 0.147{+}0.147{+}0.147{+}0.343 = 0.784$$

iii. we get at least one tail.
We can do this two ways.
1. The outcomes that cover at least one tail are HHT, HTH, HTT, THH, THT, TTH or TTT

$$\begin{aligned} P(\textbf{at least one tail}) &= P(\textbf{HHT, HTH, HTT, THH, THT, TTH or TTT}) \\ &= 0.147 + 0.147 + 0.063 + 0.147 + 0.063 + 0.063 + 0.027 \\ &= 0.657 \end{aligned}$$

2. We can use the complement rule. The complement of at least one tail is no tails.

$$\begin{aligned} P(\textbf{at least one tail}) &= 1 - P(\textbf{no tails}) \\ &= 1 - P(HHH) \\ &= 1 - .343 \\ &= 0.657 \end{aligned}$$

iv. we get a tail on the second toss.
The outcomes that cover a tail on the second coin toss are HTH, HTT, TTH, or TTT.

$$\begin{aligned} P(\textbf{tail on second toss}) &= P(\textbf{HTH, HTT, TTH, or TTT}) \\ &= (.7)(.3)(.7) + (.7)(.3)(.3) + (.3)(.3)(.7) + (.3)(.3)(.3) \\ &= .147 + .063 + .063 + .027 \\ &= .3 \end{aligned}$$

v. we get the same result on all three coin tosses.
The outcomes that give us the same result on all three coin tosses are HHH or TTT.

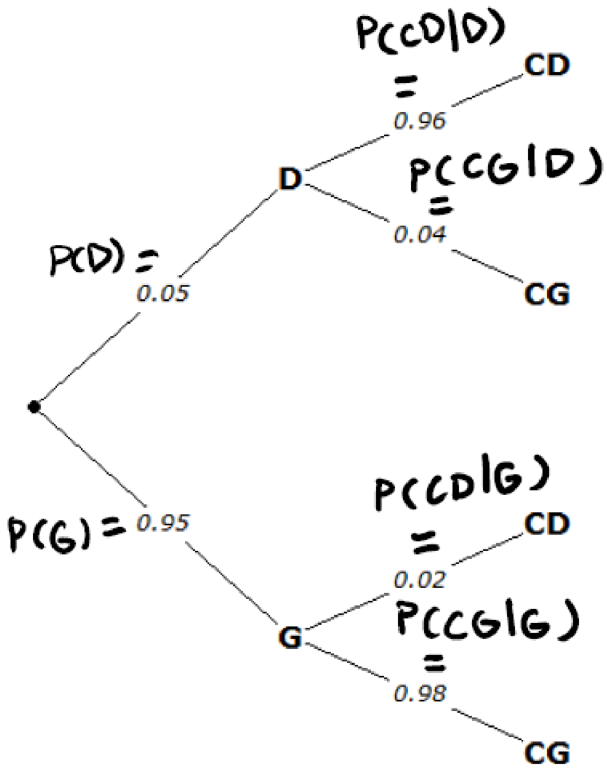
$$P(\textbf{same result on all 3 tosses}) = P(\textbf{HHH or TTT}) = 0.343 + 0.027 = 0.37$$

15. A construction company employs several welders who do thousands of welds each year. From previous records, the company knows that 5% of all welds done by their welders do not meet industry safety requirements; that is, the welds are considered defective (D). As such, each weld that is done on a construction project is examined by the company's inspector. The inspector's performance has also been monitored, and over several years the following characteristics have been observed:

- Any weld that is in fact defective (D) will be correctly classified as defective (CD) by the inspector 96% of the time; in the remaining cases, the defective weld will be classified as being good (CG);
- Any weld that is good (G) will be incorrectly classified as being defective (CD) 2% of the time; in the remaining cases, the good weld will be classified as being good (CG).

(a) Draw a tree diagram that represents all the possible outcomes associated with the experiment of selecting a weld at random for inspection and the classification of the weld by the inspector.

(b) If a randomly selected weld is classified by the inspector as being good (CG), what is the probability that the weld is defective (D)?



$$P(CG) = P(D \text{ and } CG) + P(G \text{ and } CG) = (0.05)(0.04) + (0.95)(0.98) = 0.002 + 0.931 = 0.933$$

$$P(D|CG) = \frac{P(D \text{ and } CG)}{P(CG)} = \frac{0.002}{0.933} = \mathbf{0.0021}$$

16. The probability that Shylor stops for a soda on his way to work is .23. If he stops for a soda there is a .76 chance he is late. Overall he shows up late 34% of the time. Today he is late, what is the chance that he stopped for a soda?

I'll do S for soda and L for late.
$P(S) = .23$, $P(L \mid S) = .76$, $P(L) = .34$

We want to find $P(S \mid L)$. So we want to "reverse" the conditional probability. So use Bayes' Theorem.

Bayes' Theorem

$$P(S \mid L) = \frac{P(L \mid S) \cdot P(S)}{P(L)} = \frac{(.76) \cdot (.23)}{.34} = .5141$$

Now you might not remember Bayes' Theorem, but that's okay because it is just a shortcut.

- **You might think, I need to find a conditional probability, that means divide.**

$$P(S \mid L) = \frac{P(L \cap S)}{P(L)}$$

- **But I don't know what $P(L \cap S)$ is. But I do know that to find an "and" I multiply**

$$P(L \cap S) = P(L \mid S) \cdot P(S) = .76\,(.23)$$

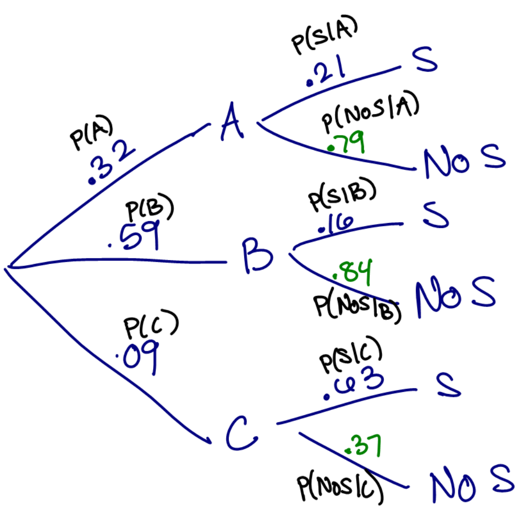
- **Now I just plug that in above.**

$$P(S \mid L) = \frac{P(L \cap S)}{P(L)} = \frac{.76\,(.23)}{.34} = .5141$$

17. Among people infected with a certain virus, 32% have strain A, 59% have strain B, and 9% have strain C.

    21% of people with strain A exhibit symptoms, 16% with strain B exhibit symptoms, 63% of people with strain C exhibit symptoms.

    (a) Draw a tree diagram representing all the possible outcomes of which strain a person has and if they exhibit symptoms.
    When I draw the tree diagrams I think that I need the first column to be the thing that happens first. So first you get infected with a strain of virus. Then you either show symptoms or you don't.



Probability of Intersections

P(A and S) = P(A)·P(S|A) = .32(.21) = .0672

P(A ∩ No S) = P(A)·P(No S|A) = .32(.79) = .2528

(.59)(.16) = .0944

(.59)(.84) = .4956

.09(.63) = .0567

.09(.37) = .0333

    (b) If they have strain B, what is the probability that they do NOT show symptoms?
    I didn't do any complicated formulas for this one.
    Instead I know that each set of branches has to add to 1. So I did $1 - .16 = 84$.

$$P(No\,S \mid B) = .84$$

(c) Find the probability that they have strain B *and* they do NOT show symptoms.
I know the formula to find the probability of an intersection. Or I can also remember that to find the probability of the intersection you multiply along the branches.

$$P(B \cap No\,S) = P(B) \cdot P(No\,S \mid B) = .59 \cdot .84 = .4956$$

(d) Find the probability that they have strain C and they show symptoms.

$$P(C \,\textbf{and}\, S) = P(C) \cdot P(S \mid C) = .09\,(.63) = .0567$$

(e) Find the overall probability that an infected person shows symptoms.
I wrote down the probabilities of all the intersections. Now add up all the probabilities that include an S for symptoms.

$$\begin{aligned} P(S) &= P(A \cap S) + P(B \cap S) + P(C \cap S) \\ &= .0672 + .0944 + .0567 \\ &= .2183 \end{aligned}$$

(f) If someone shows symptoms, find the probability that they have strain C.
To find a conditional probability, you divide.

$$P(C \mid S) = \frac{P(C \,\textbf{and}\, S)}{P(S)} = \frac{.0567}{.2183} = .2597$$

(g) Now in parts 17d and 17f we found $P(C \,\text{and}\, S)$ and then used that to find $P(C \mid S)$. Sometimes we want to combine those two steps into 1.
Using Bayes' Theorem, find $P(C \mid S)$ in one step.
$P(C) = .09$, $P(S) = .2183$, $P(S \mid C) = .63$
Bayes' Theorem

$$P(C \mid S) = \frac{P(S \mid C) \cdot P(C)}{P(S)} = \frac{.63\,(.09)}{.2183} = 0.2597$$

0.1.1 Practice Problems for Sections 1.6 and 1.7

I don't have problems written up, so you need to do these problems from the book.

Section 1.6: 1.10.33 (use tree diagram), 1.10.36a (use tree diagrams)

Section 1.7: 1.7.7, 1.7.5, 1.10.22

18. Remember Problem 14.

Suppose we toss a biased coin 3 times. This biased coin has 0.7 probability of being a head on any coin toss.

We found the sample space (all the possible outcomes) and the probabilities.

Outcome	(Show your work)	Probability
HHH	$P(HHH) = P(H) \cdot P(H) \cdot P(H) = (0.7)(0.7)(0.7) = 0.343$	0.343
HHT	$(.7)(.7)(.3)$	0.147
HTH	$(.7)(.3)(.7)$	0.147
HTT	$(.7)(.3)(.3)$	0.063
THH	$(.3)(.7)(.7)$	0.147
THT	$(.3)(.7)(.3)$	0.063
TTH	$(.3)(.3)(.7)$	0.063
TTT	$(.3)(.3)(.3)$	0.027

Let X be the number of heads in 3 tosses.

(a) Find the probability distribution (pmf) of X.

Hint: to find $P(X = 0)$ add up the probability for all the ways you can get 1 head. So add up the probabilities for HTT, THT, TTH.

So $P(X = 1) = .063 + .063 + .063$

And $P(X = 2) = .147 + .147 + .147 = .411$

x, **number of heads**	**probability of** x
0	**.027**
1	**.189**
2	**.441**
3	**.343**

****You might also realize this would be the binomial distribution with $p = .7$ and $n = 3$. Then you can use the binomial formula to find the probability of each x instead of doing it by hand.**

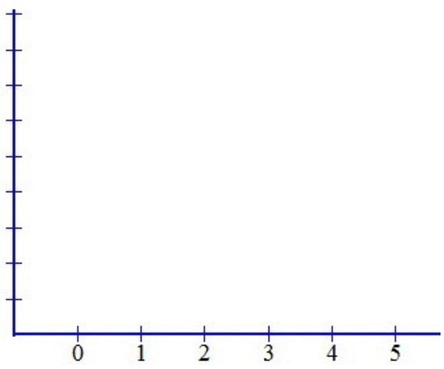
(b) Check that this is a valid probability distribution.

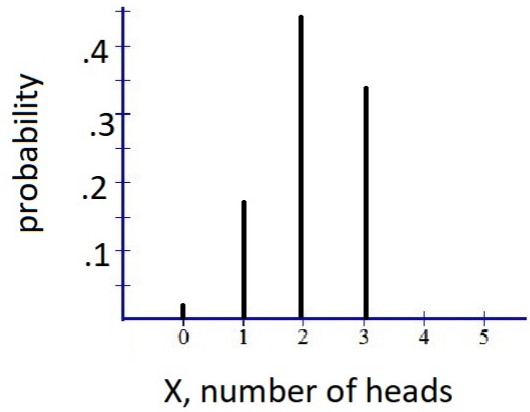
1. Are all the probabilities between 0 and 1? Yes
2. Do the probabilities add up to 1? Yes

$$.027 + .189 + .441 + .343$$

So it is a valid probability distribution.

(c) Graph the probability distribution (pmf) of X. (Don't forget to label the axes.)





(d) Find and graph the cdf of X.

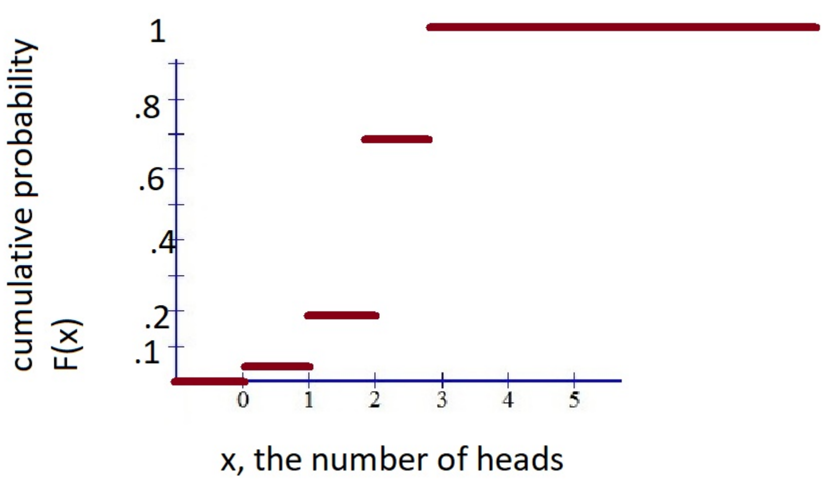
Remember that the cdf is the CUMULATIVE distribution function. So for each x value you are adding up all the probability that came before it.

x	$f(x)$	$F(x)$
number of heads	**probability of x, pmf**	**cumulative probability, cdf**
0	**.027**	**.027**
1	**.189**	**.216**
2	**.441**	**.657**
3	**.343**	**1**

Now to make it more complicated, the person who came up with the cdf says it has to be defined for all x values. So what we have above is a less technical version. Technically you should write something like what I have below.

x	$F(x)$
$x < 0$	**0**
$0 \le x < 1$	**.027**
$1 \le x < 2$	**.216**
$2 \le x < 3$	**.657**
$x \ge 3$	**1**

Because it is a discrete random variable, we only add more probability to the CUMULATIVE probability on the values $x = 0, 1, 2, 3$ so those are places where the graph jumps.



19. A basketball player believes that the probability distribution for the number of free throws he makes out of two attempts is:

Value of X	0	1	2
Probability	.36	.28	.36

(a) Calculate the mean (expected value) of the number of hits.

$$\begin{aligned} \mu_X &= x_1p_1 + x_2p_2 + \cdots + x_kp_k \\ &= 0\,(.36) + 1\,(.28) + 2\,(.36) \\ &= 1 \end{aligned}$$

So if he makes many, many, many two free throw attempts, on average he will make 1 out of the two attempts.

(b) Calculate the standard deviation σ_X of the number of hits.

$$\begin{aligned} \sigma_X &= \sqrt{(x_1-\mu_X)^2 p_1 + (x_2-\mu_X)^2 p_2 + \cdots + (x_k-\mu_X)^2 p_k} \\ &= \sqrt{(0-1)^2(.36) + (1-1)^2(.28) + (2-1)^2(.36)} \\ &= \sqrt{.72} \\ &= .85 \end{aligned}$$

20. Which variables are discrete?

(a) The number of four patients taking a new antibiotic who experience gastrointestinal distress as a side effect.
Discrete.
The only possible values are 0, 1, 2, 3, or 4.

(b) The number of the next three customers entering the store who will make a purchase.
Discrete.
The only possible values are 0,1,2, or 3.

(c) The weight of strawberry preserves dispensed by an automatic filling machine into a 16-ounce jar.
Continuous.
Assuming that you can't have negative weight and you can't fit more than 18 ounces in the jar...
The possible values are anything between 0 and 18 ounces. But most of the values are hopefully close to 16 ounces.

(d) The score on a test that is graded in half point increments.
Discrete.
The ONLY possible values are 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, ..., 29, 29.5, or 30.

(e) The number of fires in a large city in the next two months.
Discrete.
The only possible values are 0, 1, 2, 3, 4, etc.

(f) The time that a customer in a store must wait to receive a credit card authorization.
Continuous
Time is always theoretically continuous because it could be 10 minutes and 5 seconds and 1.2231 milliseconds.

21. Which of the following might be reasonably modeled by the binomial distribution?

(a) The number of customers that enter a store in a one–hour period, assuming customers enter independently.

(b) The number of questions you get correct on a 100–question multiple choice exam in which each question has four possible answers and exactly one of these answer is correct. Assume you have not studied at all for the test and have to guess each answer.

(c) A couple will have children until they have three girls or five children. X is the number of children in the family.

(d) A really nervous basketball player is making two free throw attempts. The probability of him making the first shot is .4, but if he misses the first shot, the probability of him making the second shot is only .1. X is the number of shots he makes.
(b) is binomial with $n = 100$ and $p = .25$.
(a) and (c) are NOT binomial because we don't have a fixed number of observations.
(d) is not binomial because the probability of making a shot changes. This also means that the shots are not independent.

22. The weight of medium-size tomatoes selected at random from a large bin at the local supermarket is a random variable X with mean $\mu_X = 10$oz. and standard deviation $\sigma_X = 1.2$oz.
Hint: Review section 3.6 in the course reader and rewatch the videos.
https://youtu.be/dGzdp1XUVGA
https://youtu.be/EzJEhuX-7-k

Video for this problem (but try the problem before watching the video): https://youtu.be/RrzKxG-pgFU

(a) Let the random variable W be the weight of the tomatoes in pounds (1 pound = 16 oz). What is the expected value and standard deviation σ_W of the random variable W (in pounds)?
Use the rules for means and variances. $E(X) = 10$ **and** $SD(X) = 1.2$, $Var(X) = 1.2^2$
Notice that you are given the standard deviation $\sigma_X = 1.2$, **so if you need the variance, you have to square it.**
If there are 16 oz in 1 pound, then we can just divide the weight in ounces to get to pounds.
The weight in pounds is

$$W = \frac{1}{16}X$$

So the expected value is

$$E(W) = E\left(\frac{1}{16}X\right) = \left(\frac{1}{16}\right) \cdot E(X) = \frac{1}{16} \cdot 10 = .625$$

The variance is

$$Var(W) = Var\left(\frac{1}{16}X\right) = \left(\frac{1}{16}\right)^2 \cdot Var(X) = \left(\frac{1}{16}\right)^2 \cdot (1.2)^2 = .005625$$

Now take the square root to get standard deviation.

$$SD(X) = \sqrt{.005625} = .075$$

(b) Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable Y as the weight of the content of the bag containing the four tomatoes. What is the mean μ_Y of the random variable Y (in oz.)? What is the standard deviation?

$E(X) = 10$ **and** $SD(X) = 1.2$, $Var(X) = 1.2^2$

The rules only work for variance. So find we have to work with the variance. If the standard deviation is 1.2, then the variance is 1.2^2**.**

The total weight in bag is

$$Y = X_1 + X_2 + X_3 + X_4$$

So the expected value is

$$E(Y) = E(X_1 + X_2 + X_3 + X_4) = E(X_1) + E(X_2) + E(X_3) + E(X_4) = 10 + 10 + 10 + 10 = 40$$

The variance is

$$Var(Y) = Var(X_1 + X_2 + X_3 + X_4) = Var(X_1)+Var(X_2)+Var(X_3)+Var(X_4) = 1.2^2+1.2^2+1.2^2+1.2^2 = 5.76$$

Now take the square root to get standard deviation.

$$SD(Y) = \sqrt{5.76} = 2.4$$

*****Because we are picking 4 different tomatoes, each tomato should have a different weight. So we can't just pick one tomato and multiply its weight by 4. Because again, each tomato is different. So we have to do**

$$Y = X_1 + X_2 + X_3 + X_4$$

We can NOT do $Y = 4X$**.**

(c) Suppose we pick two tomatoes at random from the bin. Let the random variable V be the difference in the weights (in oz.) of the two tomatoes selected (i.e., the weight of the first tomato minus the weight of the second tomato). What is the expected value μ_V of the random variable V (in oz.)? What is the standard deviation?

$E(X) = 10$ **and** $SD(X) = 1.2$, $Var(X) = 1.2^2$

The difference is

$$V = X_1 - X_2$$

So the expected value is

$$E(V) = E(X_1 - X_2) = E(X_1) - E(X_2) = 10 - 10 = 0\,\textbf{oz}$$

The variance is

$$Var(V) = Var(X_1 - X_2) = Var(X_1) + (-1)^2\,Var(X_2) = 1.2^2 + (-1)^2 \cdot 1.2^2 = 2.88$$

Now take the square root to get standard deviation.

$$SD(V) = \sqrt{2.88} = 1.697\,\textbf{oz}$$

****There is no formula for** $X - Y$**. You have to use a little trick.**
We remember from math class that
$5 - 3$ **is really** $5 + (-3)$**. (Subtracting is the same as adding a negative).**
So we use the trick.

$X - Y$ **is the same as** $X + (-1) \cdot Y$
Then you can use the rules we know. Like to pull the -1 out of the variance you have to square it. So

$$(-1)^2$$

23. A student takes a multiple choice test and guesses on 3 questions. Each question has **4 possibilities.**

(a) Is $X =$ the number of correct answers a binomial distribution? If so, what are n and p?

- **There are only** $n = 3$ **observations.** $\surd$
- **Each question should be independent.** $\surd$
- **Each guess is either "correct" or a "not correct".** $\surd$
- **The probability of each guess being correct is** $p = .25$. $\surd$
- **BINOMIAL** $n = 3$, $p = .25$

(b) Find the probability distribution of X.
You can use the binomial formula or your calculator.

X	**probability**
0	$\binom{3}{0}(.25)^0(.75)^3 = .4219$
1	$\binom{3}{1}(.25)^1(.75)^2 = .4219$
2	$\binom{3}{2}(.25)^2(.75)^1 = .1406$
3	$\binom{3}{3}(.25)^3(.75)^0 = .0156$

****Don't forget that** $\binom{3}{1}$ **is a binomial coefficient. It tells you how many ways you can choose 1 question out of the 3 to be the one you get correct.**
You put it in your calculator by looking for "combination" or something like "nCr" or "nCk".
Search google for your calculator model and "combination" to see how to do it.
Review page 69 in the course reader for more information.
Or read https://www.mathsisfun.com/combinatorics/combinations-permutations.html (Just the part that says Combinations without replacement)

(c) Is this a discrete or continuous distribution?
Discrete. There are only 4 possible values of X. **Also, the binomial distribution is always discrete.**

(d) Find μ_X.
Since it is binomial, you can use the formula

$$\mu_X = np = 3(.25) = .75$$

You could do also it by hand

$$\begin{aligned} \mu_X &= x_1p_1 + x_2p_2 + \cdots + x_kp_k \\ &= 0(4219) + 1(.4219) + 2(.1406) + 3(.0156) \\ &= .75 \end{aligned}$$

(e) Find σ_X.
Since it is binomial, you can use the formula

$$\sigma_X = \sqrt{np(1-p)} = \sqrt{3(.25)(.75)} = .75$$

You could also do it by hand

$$\begin{aligned}
\sigma_X &= \sqrt{(x_1-\mu_X)^2 p_1 + (x_2-\mu_X)^2 p_2 + \cdots + (x_k-\mu_X)^2 p_k} \\
&= \sqrt{(0-.75)^2(.4219) + (1-.75)^2(.4219) + (2-.75)^2(.1406) + (3-.75)^2(.0156)} \\
&= \sqrt{.5625} \\
&= .75
\end{aligned}$$

24. During the last student elections at a certain college, 45% of the students voted for the democratic student party. A simple random sample of students from this college is to be selected.

    (a) If 12 students are to be selected, what is the distribution of the number of students in the sample who voted for the democratic student party?
    This is a binomial distribution.

    - **There are only $n = 12$ observations.** $\surd$
    - **Each person should be independent.** $\surd$
    - **Each person is either democratic" or "Not democratic".** $\surd$
    - **The probability of each person being democratic is .45.** $\surd$
    - **BINOMIAL $n = 12$, $p = .45$**

    (b) What is the expected number of students in the sample that voted for the democratic student party?
    Since this is a binomial distribution we can use the formula

    $$\mu_X = np = 12\,(.45) = 5.4$$

    Of course, we can't have 5.4 students in one sample, but if we took many samples of size 12, on average each sample will have 5.4 students that voted for the democratic party.

    (c) What is the standard deviation of the random variable, σ_X?
    Since this is a binomial distribution we can use the formula

    $$\sigma_X = \sqrt{np\,(1-p)} = \sqrt{12\,(.45)\,(1-.45)} = \sqrt{2.97} = 1.7234$$

    (d) What is the variance of the random variable, σ_X^2?
    The variance is always the square of the standard deviation.

    $$\sigma_X^2 = (1.7234)^2 = 2.97$$

    (e) If 12 students are to be selected, what is the probability that more than 7 students in the sample voted for the democratic student party?
    Use the binomial formula or your calculator.

    $$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\begin{aligned}
P(X>7) &= P(X=8) + P(X=9) + P(X=10) + P(X=11) + P(X=12) \\
&= \binom{12}{8}(.45)^8(1-.45)^{12-8} + \binom{12}{9}(.45)^9(1-.45)^{12-9} + \cdots \\
&= \binom{12}{8}(.45)^8(.55)^4 + \binom{12}{9}(.45)^9(.55)^3 + \cdots \\
&= .0762 + .0277 + .0068 + .00101 + .0000689 \\
&= .1118
\end{aligned}$$

25. An insurance company will insure a $75,000 Hummer for its full value against theft at a premium of $1500 per year. Suppose that the probability that the Hummer will be stolen is 0.0075.

    (a) Calculate the insurance company's expected net profit for one year.

X (profit)	probability
1500	**.9925**
-73500	**.0075**

    $$\begin{aligned} \mu_X &= 1500\,(.9925) + (-73500)\,(.0075) \\ &= 937.50 \end{aligned}$$

    (b) Interpret your results.
    If the company sells many, many Hummer policies, they will make an average of $937.50 per policy per year.

26. Universal blood donors: People with type O–negative blood are universal donors. That is, any patient can receive a transfusion of O–negative blood. About 6% of a particular population have O–negative blood. If 12 people appear at random to give blood:
This is a binomial distribution.

    - **There are only** $n = 12$ **observations.** $\surd$
    - **Each person should be independent.** $\surd$
    - **Each person is either "Type O-" or "Not Type O-".** $\surd$
    - **The probability of each person being type O-negative is** $p = .06$**.** $\surd$
    - **BINOMIAL** $n = 12$**,** $p = .06$

    (a) What is the probability that at least 1 of them is a universal donor?
    You can do this without knowing that it is the binomial distribution. The probability that any one person is not O-negative is $.94$**.**

    $$\begin{aligned} P\,(\textbf{at least 1 type Oneg}) &= 1 - P\,(\textbf{no O- blood}) \\ &= 1 - P\,(12\,\textbf{people not blood}) \\ &= 1 - (.94)^{12} \\ &= 1 - .4759 \\ &= .5241 \end{aligned}$$

    You can also do it by using the binomial formula.

    $$\begin{aligned} P\,(X \geq 1) &= P\,(X = 1) + P\,(X = 2) + P\,(X = 3) + \cdots + P\,(X = 12) \\ &= .3645 + .1280 + .0272 + .0039 + .0004 + \mathbf{0} \\ &= .5240 \end{aligned}$$

    Or you can use binomial formula and the complement rule.

    $$\begin{aligned} P\,(X \geq 1) &= 1 - P\,(X = 0) \\ &= 1 - \binom{12}{0}(.06)^0\,(.94)^{12} \\ &= 1 - .4759 \\ &= .5241 \end{aligned}$$

    (b) What is the probability that 3 of them are universal donors?
    You can do it by using the calculator:

    $$P\,(X = 3) = .0272$$

You could do it by hand if you want.

$$P(\mathbf{X{=}3}) = \binom{12}{3}(.06)^3(1-.06)^9 = .0272$$

(c) What is the probability that there are 3 or fewer universal donors?
3 or fewer means it could be 0 or 1 or 2 or 3.
You can find each one by hand or using your calculator and then add them up.

$$\begin{aligned} P(X \le 3) &= P(X=0) + P(X=1) + P(X=2) + P(X=3) \\ &= .4759 + .3645 + .1280 + .0272 \\ &= .9959 \end{aligned}$$

If you really want you could use the binomial cdf function on your calculator if you know how to find $P(X \le 3)$ and get the answer in one step.

27. Portfolio analysis. Here are the means and standard deviations for the annual returns from three Fidelity mutual funds for the 10 years ending in February 2004. Assume that the stocks returns are independent of each other.

W =annual return on 500 Index Fund
X =annual return on Investment Grade Bond Fund
Y =annual return on Diversified International Fund
$\mu_W = 10.12\%$, $\sigma_W = 17.46\%$
$\mu_X = 6.46\%$, $\sigma_X = 4.18\%$
$\mu_Y = 11.10\%$, $\sigma_Y = 15.62\%$

Assume that a portfolio contains 70% Investment Grade Bond Fund (X) and 30% Diversified International Fund (Y) stocks. Calculate the mean and standard deviation of the returns for this portfolio.
This comes from Section 3.6 Combinations of Random Variable

Review from lecture: https://youtu.be/dGzdp1XUVGA

Video for this problem: https://youtu.be/4TqaR7hSzeY

$R = .7X + .3Y$

$$E(R) = E(.7X + .3Y) = .7 \cdot E(X) + .3 \cdot E(Y) = .7(6.46) + .3(11.10) = 7.852\%$$

$$\begin{aligned} Var(R) &= Var(.7X + .3Y) \\ &= (.7)^2 Var(X) + (.3)^2 Var(Y) \\ &= (.7)^2 (4.18)^2 + (.3)^2 (15.62)^2 \\ &= 30.52 \end{aligned}$$

$$SD(Y) = \sqrt{30.52} = 5.52\%$$

28. A variable is called random if

(a) Individual outcomes are uncertain but happen in a predictable manner through time.

(b) One has no idea what will happen.

(c) Outcomes happen in a 50–50 split.

(d) I flip a coin and get 5 heads on my first 5 flips, I am very likely to get 5 tails on my next 5 flips.
(a) For example, when I flip a coin, I don't know if I will get a head or a tail; but I do know that if I flip the coin many times, about 50% of the tosses will result in heads.

29. Mandy usually gets paid \$5 from customers on her newspaper route. One day a customer offers Mandy a deal. He will place three \$1 bills, one \$5 bill, and one \$10 bill in a paper bag. Mandy can then draw and keep one bill from the bag instead of her usual \$5. She could do this every week from now on.

(a) **Gut feeling:** Do you think Mandy should take the deal? Will she make more or less money in the long run?
No correct answer. It is your gut feeling.

(b) Find the expected value of Mandy's earnings if she takes the deal.
First, you need to figure out the possibilities of Mandy's earnings and the probabilities of each possibility.

Mandy's Earnings	Probability
\$1	3/5
\$5	1/5
\$10	1/5

The expected value is

$$\mu_X = E(X) = 1 \cdot \frac{3}{5} + 5 \cdot \frac{1}{5} + 10 \cdot \frac{1}{5} = 3.6$$

(c) Interpret your result.
In the long run, over many weeks, Mandy will make an average of \$3.6 per week.

(d) Will Mandy make more or less money in the long run if she takes the deal?
Mandy will only make on average \$3.6 per week instead of \$5 every week. So she will lose money in the long run.

30. Remember the sentences given by a second judge have a distribution given by

$$f(y) = A\sqrt{y}, \qquad 0 \le y \le 4$$

(a) What must A be in order for this to be a pdf? **This is example 68 and 69 in the guided notes https://youtu.be/fswe8kM-300**

Example 68. Suppose the sentences given by a second judge have a different distribution given by

$$f(y) = A\sqrt{y}, \qquad 0 \le y \le 4$$

constant

What must A be in order for this to be a pdf?

$\int f(x)dx = 1$

$\int_0^4 A\sqrt{y}\,dy = \frac{16}{3}A$

supposed to be 1

$\frac{16}{3}A = 1$

$A = 3/16$

so $f(y) = \frac{3}{16}\sqrt{y}$

(b) Find the cdf.

Example 69. Remember the sentences given by a second judge have a distribution given by

$$f(y) = \frac{3}{16}\sqrt{y}, \qquad 0 \le y \le 4 \quad \text{pdf}$$

1. Find the cdf. pdf → cdf

$F(y) = \int_{-\infty}^{y} f(t)dt = \int_0^y \frac{3}{16}\sqrt{t}\,dt = \frac{1}{8}y^{3/2}$

$\boxed{F(y) = \frac{1}{8}y^{3/2} \text{ for } 0 \le y \le 4}$

2. Use the cdf to find the probability that a prison term is no more than 2.5 years. cdf → probability

$P(0 \le X \le 2.5) = F(2.5) - F(0)$

$= \left[\frac{1}{8}(2.5)^{3/2}\right] - \left(\frac{1}{8}(0)^{3/2}\right) = \boxed{.4941}$

3. Use the pdf to find the probability that a prison term is at least 2.5 years. pdf → probability

$P(2.5 \le X \le 4) = \int f(y)dy = \int_{2.5}^{4} \frac{3}{16}\sqrt{y}\,dy = \boxed{.5059}$

4. Use the cdf to find the probability that a prison term is more than 2.5 years. cdf → probability

$P(2.5 \le X \le 4) = F(4) - F(2.5)$

$= \frac{1}{8}(4)^{3/2} - \frac{1}{8}(2.5)^{3/2} = \boxed{.5059}$

(c) Use the cdf to find the probability that a prison term is no more than 2.5 years.
****Note that each probability can be found with either the cdf or pdf.**
We want to find $P(Y \le 2.5)$.
Method 1: cdf, plug the y value into the cdf.
The cdf is $F(y) = \frac{1}{8}y^{3/2}$ for $0 \le y \le 4$.

$$P(Y \le 2.5) = F(2.5) = \frac{1}{8}(2.5)^{3/2} = .4941$$

****

Method 2: pdf, integrate the pdf with the bounds 0 to 2.5.
The pdf is $\frac{3}{16}y^{1/2}$ for $0 \le y \le 4$.

$$P(Y \le 2.5) = \int_0^{2.5} \frac{3}{16}y^{1/2}dy = .4941$$

(d) Use the pdf to find the probability that a prison term is at least 2.5 years.
The pdf is $\frac{3}{16}y^{1/2}$ **for** $0 \leq y \leq 4$**.**
We want to find $P(Y \geq 2.5)$**.**
pdf, integrate the pdf with the bounds 2.5 to the upper limit of 4.

$$P(Y \geq 2.5) = \int_{2.5}^{4} \frac{3}{16}y^{1/2}dy = .5059$$

(e) Use the cdf to find the probability that a prison term is more than 2.5 years. **The cdf is** $F(y) = \frac{1}{8}y^{3/2}$ **for** $0 \leq y \leq 4$**.**
We want to find $P(Y \geq 2.5)$**.**
Use the cdf:
But the problem is that we want the probability $Y \geq 2.5$ **and cdf is built to give us the probability** $Y \leq 2.5$**. So you have to use the complement rule.**

$$P(Y \geq 2.5) = 1 - F(2.5) = 1 - \frac{1}{8}(2.5)^{3/2} = .5059$$

****Notice that since** $y = 4$ **is the upper bound, you could also do** $P(X \geq 2.5)$ **is the same as** $P(2.5 \leq Y \leq 4)$

$$\begin{aligned} P(2.5 \leq Y \leq 4) &= F(4) - F(2.5) \\ &= \frac{1}{8}(4)^{3/2} - \frac{1}{8}(2.5)^{3/2} \\ &= 1 - \frac{1}{8}(2.5)^{3/2} \\ &= .5059 \end{aligned}$$

31. A cdf is

$$F(y) = \frac{y^{3/2}}{8}, \qquad 0 \le y \le 4$$

(a) Use the cdf to find $P(-2 \le Y < 3.6)$.

(b) Find the probability density function.

(c) Use the pdf to find the probability that Y is between 1.2 and 3.

Video: Section 2.2 Example for cdf and pdf

$F(y) = P(Y \le y)$

Example 70. A cdf is

$$F(y) = \frac{y^{3/2}}{8}, \qquad 0 \le y \le 4$$

1. Use the cdf to find $P(-2 \le Y < 3.6)$. this is really $P(0 \le Y < 3.6)$

$F(3.6) - F(0)$

$= \frac{3.6^{3/2}}{8} - \frac{0^{3/2}}{8} = .8538$

2. Find the probability density function. cdf $\rightarrow$ pdf take derivative

pdf $f(y) = \frac{d}{dy} F(y) = \frac{d}{dy}\left(\frac{y^{3/2}}{8}\right) = \frac{1}{8} \cdot y^{1/2}\,(3/2) = \frac{3}{16} \cdot y^{1/2}$

$f(y) = \frac{3}{16}\sqrt{y}$ for $0 \le y \le 4$

3. Use the pdf to find the probability that $\cancel{X}$ Y is between 1.2 and 3.

$P(1.2 \le Y \le 3) = \int_{1.2}^{3} \frac{3}{16}(y)^{1/2}\,dy = \frac{3}{16} \cdot \frac{y^{3/2}}{3/2} \Big|_{1.2}^{3}$

$= \frac{1}{8} y^{3/2} \Big|_{y=1.2}^{3} = \left(\frac{1}{8} \cdot 3^{3/2}\right) - \left(\frac{1}{8} \cdot (1.2)^{3/2}\right) = .4852$

32. Remember the sentences given by a second judge have a distribution given by

$$f(y) = \frac{3}{16}\sqrt{y}, \qquad 0 \le y \le 4$$

Find the standard deviation of the length of the sentences.

Example 77. Remember the sentences given by a second judge have a distribution given by

$$f(y) = \frac{3}{16}\sqrt{y}, \qquad 0 \le y \le 4$$

Find the standard deviation of the length of the sentences.

$$E(y) = \mu = \int y \cdot f(y)\,dy = \int_0^4 y \cdot \frac{3}{16}\sqrt{y}\,dy = 2.4$$

$$Var(Y) = \int (y-\mu)^2 \cdot f(y)\,dy = \int_0^4 (y-2.4)^2 \cdot \frac{3}{16}\sqrt{y}\,dy = 1.097$$

$$\sigma = \sqrt{1.097} = \boxed{1.047}$$

2nd formula $Var(Y) = E(Y^2) - [E(Y)]^2$

$$E(Y) = \int y \cdot f(y)\,dy = \int_0^4 y \cdot (3/16\sqrt{y})\,dy = 2.4$$

$$E(Y^2) = \int y^2 \cdot f(y)\,dy = \int_0^4 y^2 (3/16\sqrt{y})\,dy = 6.85$$

$$Var(Y) = E(Y^2) - [E(Y)]^2 = 6.85 - (2.4)^2 = \boxed{1.09}$$

33. A pdf is

$$f(x) = \frac{3}{1000}x^2$$

for $0 \le x \le 10$ and $f(x) = 0$ otherwise.

(a) Find the cdf.

(b) What is the median?

(c) Find the .37 quantile.

Example 78. A pdf is

$$f(x) = \frac{3}{1000}x^2$$

for $0 \le x \le 10$ and $f(x) = 0$ otherwise.

1. Find the cdf.

$$F(x) = \int_0^x \frac{3}{1000} t^2 \, dt = \frac{3}{1000} \cdot \frac{t^3}{3} \Big|_0^x = \frac{x^3}{1000} - \frac{0^3}{1000} = \frac{x^3}{1000}$$

for $0 \le x \le 10$

2. What is the median? find $F(x) = .5$

$\frac{x^3}{1000} = .5$ solve for x

$x = 7.937$ median

3. Find the .37 quantile. find $F(x) = .37$

$\frac{x^3}{1000} = .37$ solve for x, so $x = 7.179$

.37 quantile is 7.179

37^{th} percentile is 7.179

34. We have a random variable X with pdf

$$f(x) = 3x^2 \qquad \text{for } 0 \le x \le 1$$

and $Y = X^3$. Find $E(4Y + 2)$.

Example 85. We have a random variable X with pdf

$$f(x) = 3x^2 \qquad \text{for } 0 \le x \le 1$$

and $Y = X^3$. Find $E(4Y + 2)$.

$$E(4Y+2) = E(4 \cdot X^3 + 2) = \int (4x^3+2) \cdot f(x)\,dx = \int_0^1 (4x^3+2)(3x^2)\,dx$$

$$= \boxed{4}$$

35. We have a random variable X with pdf

$$f_X(x) = 2x \qquad 0 \le x \le 1$$

and $Y = \sqrt{X}$.

Example 87. We have a random variable X with pdf

$$f_X(x) = 2x \qquad 0 \le x \le 1$$

↑ this is the pdf of X

and $Y = \sqrt{X}$.

1. Find the cdf of X. pdf to cdf

$$F_X(x) = \int_{-\infty}^{x} f(x)dx = \int_0^x 2t\,dt = x^2$$

$$\boxed{F_X(x) = x^2 \text{ for } 0 \le x \le 1}$$

cdf $F_X(x) = P(X \le x)$

2. Find the cdf of Y. trick for changing variables

$$F_Y(y) = P(Y \le y) \underset{\text{substitute } Y=\sqrt{X}}{=} P(\sqrt{X} \le y)$$

↑ solve inequality for X

$$= \underbrace{P(X \le y^2)}_{\text{cdf of } X \text{ evaluated at } x = y^2} = (y^2)^2 = y^4$$

plug y^2 into $F_X(x)$

change limits
if $x=0$, $y=\sqrt{x}$ or $\sqrt{0}=0$
if $y=1$, then $y=\sqrt{1}=1$
$F_Y(y) = y^4$
for $0 \le y \le 1$

3. Find the pdf of Y. cdf → pdf derivative

$$\frac{d}{dy}F_Y(y) = \frac{d}{dy}(y^4) = 4y^3$$

$$\boxed{f_Y(y) = 4y^3 \text{ for } 0 \le y \le 1}$$

4. Find $E(Y)$.

$$E(Y) = \int y \cdot f_Y(y)dy = \int_0^1 y \cdot (4y^3)dy = \boxed{.8}$$

5. Find $P(Y > .3)$.

$$\text{cdf: } P(Y > .3) = 1 - \underbrace{P(Y \le .3)}_{\text{cdf}} = 1 - F_Y(.3) = 1 - (.3)^4 = \boxed{.9919}$$

$$\text{pdf: } P(Y > .3) = P(.3 < Y \le 1) = \int_{.3}^{1} f_Y(y)dy = \int_{.3}^{1} 4y^3 dy = \boxed{.9919}$$

0.2 Book Problems

Here are more practice problems from the book.

Section 1.10: 1, 4, 5, 6, 7, 10, 14, 17, 18, 22, 23, 25, 26, 28, 30, 32, 33, 34, 35, 36, 37

Problem 1.10.32 has a typo in the book solutions. It should be $1 - (.0005 \times .001) = .9999995$

Section 2.9: 2, 4, 5, 7, 11, 12, *14, 15, 16, 17, 18, 20, 21, 22, 23, 26, 28

Section 3.8: 1, 2, 6, 7a, 10, 13ab

Skip sections 3.4 (Poisson), 4.1 (uniform), 4.2 (exponential),

*Do the book problems last