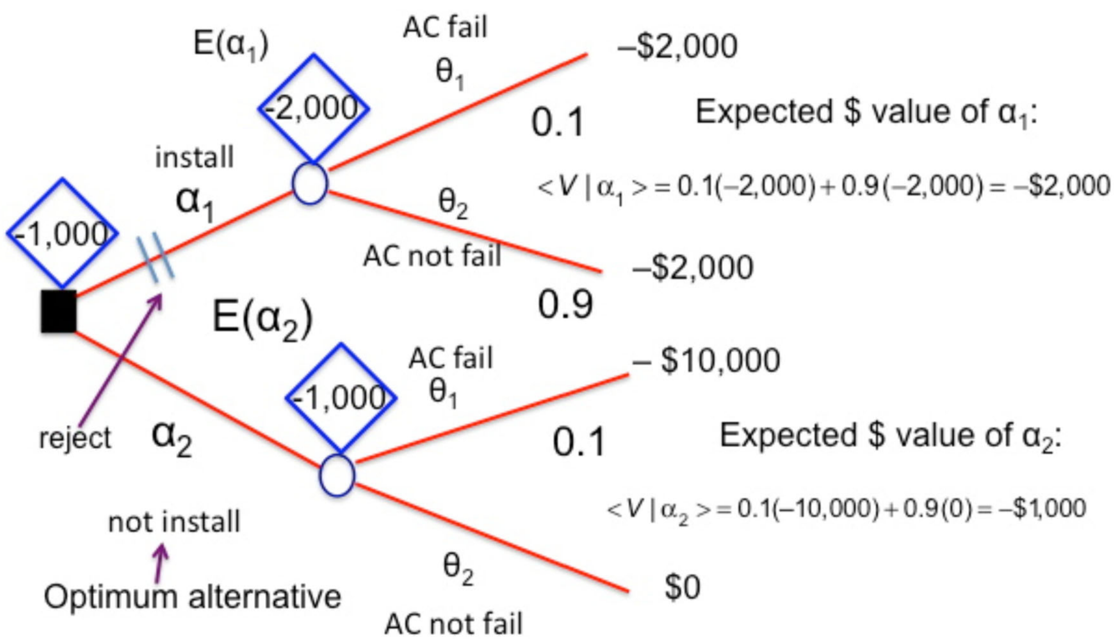
HW 1 Solution
Due Date: 02/22/2022 (11:59 PM)

1. You are to decide whether or not to install an emergency power source (EPS) to back up a part of your system in case of normal utility power failure.
   The cost of the EPS is \$2 K (hardware and maintenance for 1 year) and the approximate cost of power outage with the EPS = \$0. The cost of a power outage without the EPS = \$10 K. Based on past outages, the probability of power failure during a year is given as P(AC failure) = 0.1.
   a. Draw a decision tree based on the information provided. Show:
      i. the alternatives,
      ii. the values (or costs) of the credible outcomes given a decision,
      iii. the probabilities of each outcome.
   b. Calculate the expected monetary value (EMV) for each decision alternative and place EMV values on the decision tree. [Remember: EMV = ΣP(i)Cost(i)]
   c. Based on this information, state the optimum alternative based on the EMV of the two alternatives, and place the optimum EMV above the decision node on the tree.
   d. State whether your team, as the decision maker, would make this decision given only this information, or your team would consider other information or alternatives prior to making this decision. State examples of other considerations, or information that could be incorporated into the decision process.

Solution:

Based on the EMV amounts, the cost of Alternative II is lower, so the value is greater. Therefore, based on this information, Alternative II not to install the EPS is the optimum decision.
One consideration is indirect cost or costs of a power failure that could be added to the \$10,000 direct costs and could result in a greater expected value (lower expected costs) for Alternative I to install the EPS. Such additional costs can be incurred as a result of customer and stakeholder objections to the frequency of power outages. Also, a lower willingness to accept risk of a power outage (greater risk aversion) and other consequences of outages could favor Alternative I to install the EPS.



2. Consider the wildcatter problem that was shown in class:

a) Calculate the $P(\theta_1)$ for which the EMV of the two decisions become equivalent. Below this probability value, the EMV of alternative α_2 , not to drill, would be larger, and therefore alternative α_2 would be favored based on this information.

Solution: By Trial and Error the probability value of equal EMV is estimated to be $P(\theta_1)$~0.22 as shown below:

$$< V \mid \alpha_1 > = 0.22(10) + 0.78(-1.5) = \$1.0\,M$$

$$< V \mid \alpha_2 > = 1.0(1\ M) = \$1.0\ M$$

(b) State whether the oil drilling company should make the decision between these two alternatives based only on this information. State what other information could lower uncertainty to help them make a better decision.

Solution: The oil drilling team should consider additional information or a test of the rock strata that could lower the uncertainty of oil presence or expected quantity at the site. The value of the information (VOI) must be sufficient to offset the cost of the information. We will discuss how the VOI can be calculated and the test considered as one of the decision alternatives so the optimum alternative can be selected from alternatives, without a test and with a test, based on maximum expected utility.

(c) How would this decision depend on company size or resources. Consider this decision for two company sizes: a small company, where \$1.5 Million is a large fraction of their resources compared to a larger company with ten times the resources of the smaller company.

Solution: The decision to drill involves a relatively high probability of losing 1.5 M in drilling costs, which a larger company could absorb much more easily than a small company. Therefore compared to the large company, the smaller company would be more risk averse for this amount of monetary risk and should calculate the VOI to reduce uncertainty of oil presence prior to making a decision.

(d) If a test of the rock strata is performed and the probability of significant oil presence based on this information is increased to 0.5, calculate the expected value of the alternative to drill. State whether this monetary value should (or would likely not) offset the nominal cost of obtaining the additional information for an optimum decision.

Solution:

$$\left\langle V \mid \alpha_1 \right\rangle = 0.5(10) + 0.5(-1.5) = \$4.25\,\mathrm{M}$$

This increased monetary value of the decision alternative should easily offset the nominal cost of obtaining the additional information, such as \$50,000 for an optimum decision.

(e) If the test of the rock strata costs \$50,000 (TC), estimate the monetary value of information (VI) gained from the test, where VI = (optimum value following the test) – (optimum value without the test). To be cost effective, the VI should be significantly greater than the cost of the test.

Solution: The optimum decision alternative value without the test was \$1.375 M. With the test, the optimum value is \$4.25 M.

VI = (4.25 M) – (1.375 M) = \$2.875 M, which is significantly greater than the cost of the test. Therefore, the test is cost effective.

3. A school population consists of 35% juniors (J) and 20% seniors (S).

   a. State the general expression for the probability of J OR S, P(J∪S), which involves the addition of probability of the 2 events, J, S.

   P(J∪S) =

   b. Calculate each term in the general expression and calculate the probability (or fraction as an estimation of probability) of the student population that are juniors and seniors.

c. State whether this is an example of intersecting sets or mutually exclusive sets.

Solution:

a. General expression: $P(J \cup S) = P(J) + P(S) - P(J \cap S)$
b. Here $P(J)=0.35$, A is the set of junior students, $P(S) = 0.20$, B is the set of senior students and $P(A \cap B) = 0$, because one student cannot be junior and senior at the same time.

   So, $P(J \cup S) = 0.35+0.20-0 = 0.55$

c. Mutually exclusive sets or events, because juniors and seniors do not intersect, so they are disjoint. A student is either a junior or a senior, but not both.

4. A test was performed on 200 identical units to measure their length and height. Given that L = event that a unit fails the length criteria, H= event that a unit fails the height criteria.

   23 parts failed only the length requirement but met the height criteria. 18 parts failed only the height requirement but met the length criteria. 7 parts failed both length and height. The remaining 152 parts met the specified length and height criteria.

   a. Calculate: $P(L)$, $P(H)$, $P(L \cap H)$, $P(L \cup H)$, $P(L|H)$, $P(H|L)$
   b. Explain whether L and H are mutually exclusive.
   c. Explain whether L and H are independent.

Solution:

a. Calculate: $P(L)$, $P(H)$, $P(L \cap H)$, $P(L \cup H)$, $P(L|H)$, $P(H|L)$

P(L)=Probability of failing the length criteria = $(23+7)/200 = 30/200$

P(H)=Probability of failing the height criteria = $(18+7)/200 = 25/200$

$P(L \cap H) = 7/200$

$P(L \cup H) = P(L) + P(H) - P(L \cap H) = (30+25-7)/200 = 48/200$

$P(L|H) = P(L \cap H) / P(H) = 7/25$

$P(H|L) = 7/30$

b. Explain whether L and H are mutually exclusive.

No, they are not mutually exclusive because they can occur together. For 7 units, there is failure of both the length and height criteria.

c. Explain whether L and H are independent.

No. If L and H were independent, $P(L|H) = P(L)$ and $P(H|L) = P(H)$. However, from a., we can see that they are not so.

5. A medical study compared the success rates of two treatments for kidney stones. Each treatment was applied to two groups of people – one group in which each subject had a small stone and one group in which each subject had a large stone. Based on various numbers of patients in each case, the 'average' success rates were:

	Small Stones	Large Stones	Average
Treatment A	93%	73%	?
Treatment B	87%	69%	?

a) Fill in the table above by calculating the average value of the success rate for each treatment given the above information.

The table below shows the patient data, with the actual number of successful cases among a total of successful and unsuccessful cases for each treatment.

Average	Small Stones		Large Stones		Average
	No. of Successful Cases	Total Cases	No. of Successful Cases	Total Cases	
Treatment A	81	87	192	263	?
Treatment B	234	270	55	80	?

b) Using an overall 'Systems' approach to calculate all data together, fill in the table above by estimating point probabilities of success.
c) Compare the values in the last column of each table and explain why they are different based on how they were calculated. Based on this data, is Treatment A superior to Treatment B? Why or why not? State your reasons.
d) State briefly what is needed or what should be done to resolve with more confidence the question of effectiveness of Treatments A and B for small stones and for large stones.

Solution:

a.

	Small Stones	Large Stones	Average
Treatment A	93%	73%	**83%**
Treatment B	87%	69%	**78%**

b.

Average	Small Stones		Large Stones		Average
	No. of Successful Cases	Total Cases	No. of Successful Cases	Total Cases	
Treatment A	81	87	192	263	$\frac{\mathbf{81+192}}{\mathbf{87+263}} = \mathbf{0.78}\ (\mathbf{78}\%)$
Treatment B	234	270	55	80	$\frac{\mathbf{234+55}}{\mathbf{270+80}} = \mathbf{0.83}\ (\mathbf{83}\%)$

c. The percentages in the 4th column of first table were calculated by averaging the average values of Treatment A and B. Equal importance/ weighing was given to the data of both treatment methods. The second System approach considered more information. It put weightage on the number of cases and the subsequent successes for each treatment method. When all data are considered, treatment B appears to be superior to treatment A.

d. In general, more data will increase the credibility of an assessment. Number of cases of small stones in Treatment A and large stones in Treatment B are different. There should be a closer balance between the amount of data for each stone size for each treatment to ensure a less biased assessment.

6. Simplify the following Boolean functions:

   a) $\overline{\overline{A \cap B \cup C} \cap \overline{B}}$

   Begin with de Morgan's Theorem and then use the Law of Absorption to simplify to an expression with just B, C.

   b) $A \cap B \cap \overline{(C \cup (\bar{C} \cup A) \cup \bar{B})}$

   c) $(A \cup B) \cap \bar{B})$

   Begin with de Morgan's Theorem and then use the Commutative and Complementation laws to simplify to obtain the null set.

Solution:

a.

$\overline{\overline{(A \cap B \cup C)} \cap \overline{B}}$	de Morgan's Theorem
$A \cap B \cup C \cup B$	Law of Absorption
	$(A \cap B \cup B = B)$
$B \cup C$	

b.

$$A \cap B \cap \overline{\left[C \cup (\overline{C} \cup A) \cup \overline{B} \right]}$$

$$A \cap B \cap \left[\overline{C} \cap (C \cap \overline{A}) \cap B\right] \qquad \text{de Morgan's Theorem}$$

$$A \cap B \cap \left[(\overline{C} \cap C) \cap (\overline{A} \cap B)\right] \qquad \text{Commutative Law}$$

$$A \cap B \cap \left[(\phi) \cap (\overline{A} \cap B)\right] \qquad (\overline{C} \cap C) = (C \cap \overline{C})$$

$$(A \cap B) \cap [\phi] \qquad \text{Complementation Law} \quad (C \cap \overline{C}) = \phi$$

$$\phi$$

c.

$$(A \cup B) \cap \overline{B}) \qquad \text{Distribution Law}$$

$$(A \cap \overline{B}) \cup (B \cap \overline{B}) \qquad \text{Complementation Law}$$

$$(A \cap \overline{B}) \cup (\varnothing) \qquad \text{Union with null set}$$

$$(A \cap \overline{B})$$

7. The joint pmf (probability mass function) of precipitation, X in. and runoff, Y cfs (cubic ft/sec) is P(x,y). Because of storms at a particular location, values of the discretized joint P(x,y) are:

	X = 1	X = 2	X = 3
Y = 10	0.05	0.15	0.0
Y = 20	0.10	0.25	0.25
Y = 30	0.0	0.10	0.10

a) Write the logic expression for and calculate the probability that the next storm will result in a precipitation of 2 or more inches and a runoff of more than 20 cfs.
Hint: You will find this cumulative probabilities for which $P(X \geq 2, Y > 20)$.

Note that you will find this cumulative probability by summing the P(x,y) for $P(X \geq 2, Y > 20)$, so sum over the last row, $Y > 20$, 2nd and 3rd columns, $X \geq 2$, of the joint pmf table.

$$P(X \geq 2, Y > 20) = 0.1 + 0.1 = 0.2$$

b) Following a storm, a rain gauge shows a precipitation of 2 in. Write the logic expression for and calculate the probability that the runoff from this storm is 20 cfs or greater, $P(Y \geq 20 | X = 2)$.

$P\left(Y \geq 20 \mid X = 2\right) = \frac{P(Y \cap X)}{P(X = 2)}$, where the probability of intersection of Y and X is the sum of the two values of the X = 2 column for $Y \geq 20$, and the total probability of X = 2, $P(X = 2)$ is the sum of the values in the X = 2 column. Note that given X = 2, the sample space has been reduced from 1 (sum of all X,Y values) to 0.5.

$$P\left(Y \geq 20 \mid X = 2\right) = \frac{P(Y \cap X)}{P(X = 2)} = \frac{0.25 + 0.1}{0.15 + 0.25 + 0.10} = \frac{0.35}{0.5} = 0.7$$

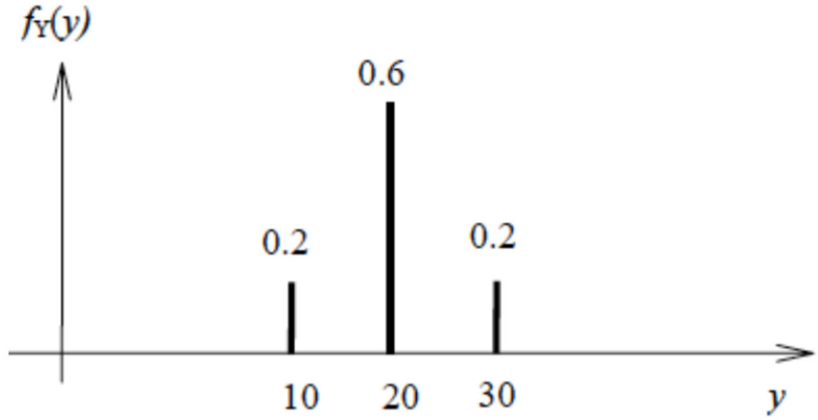
c) Check whether X and Y are statistically independent. Show why independent or why not by testing whether $P(Y \geq 20) = P(Y \geq 20|X = 2)$.

$$P(Y \geq 20) = 0.10 + 0.25 + 0.25 + 0.0 + 0.10 + 0.10 = 0.8$$

$P(Y \geq 20|X = 2) = 0.7$ as shown in b)

Because these two probability values are not identical, X and Y and not (statistically) independent.

d) By summing over each row of values in the table, write an expression for and calculate the unconditional marginal pmf of runoff, P(y). Include a sketch of the pmf, in which the 3 values of runoff Y are along the horizontal axis and P(y) or f(y) is along the vertical axis.

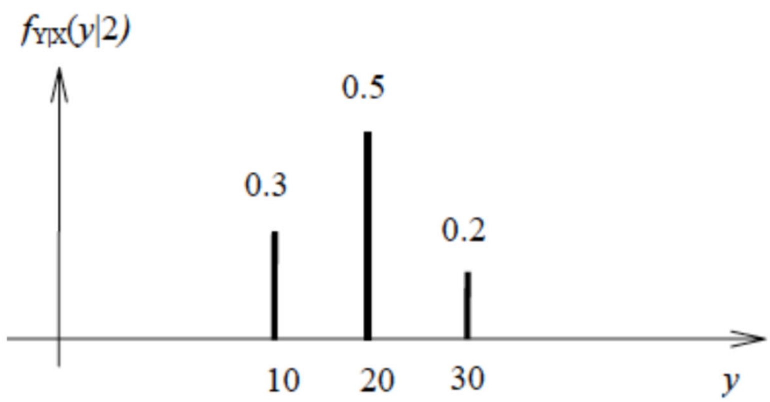


e) Write an expression for, calculate and sketch the conditional marginal pmf of runoff given a storm with 2 in precipitation is observed, [that is X = 2 in, or, P(y|x = 2)]. Check that the sum of probability values in your sketch totals 1 to satisfy Pr Axiom 2.

Given X = 2, we can use Pr Axiom 4 expression to calculate the conditional probability, P(Y|X = 2) for each of the 3 values of Y, such as

$$P\left(Y=0 \mid X=2\right)=\frac{P(Y\cap X)}{P(X=2)}=\frac{0.15}{0.15+0.25+0.10}=\frac{0.15}{0.5}=0.3$$

Note that because of the reduction in sample space given X = 2, each probability value in the X = 2 column is effectively multiplied by 2 in the conditional probability calculation based on Pr Axiom 4. So the conditional pmf for X = 2 is shown in the figure below.



8. a. The probability of failure p of a resistor is 0.08. Assuming binomial distribution, what is the probability of finding more than 1 defective resistors among a sample of 10 resistors?
   b. Because $p < 0.1$, use the Poisson distribution to calculate the probability of finding more than 1 defective resistor among 10 with acceptable agreement with the binomial result. Begin by writing an expression for ρ.
   Hint: While 'at least 1' defective resistor would mean 1 or more than 1 defective resistor, 'more than 1' would not include probability of 1 defective resistor. 'More than 1' refers to finding 2 or 3, or 4 and so on number of defective resistors.

Given the binomial distribution, where p = 0.08, q = 1–p = 0.92, and n = 10, calculate the probability of more than 1 defective resistors among the sample of 10 resistors. Begin by stating a probability expression.

$$Pr(x) = 1 - Pr(0) - Pr(1)$$

$$Pr(x) = 1 - 1(0.08)^0 (0.92)^{10} - 10(0.08)^1 (0.92)^9$$

$$Pr(x) = 1 - 0.434388 - 0.377729$$

So $Pr(x > 1) = 1 - [Pr(X = 0) + Pr(X = 1)] = 0.188 \sim 0.19$

Sketches of the binomial pmf for this problem will look similar to the figure from U8, Slide 26, where the binomial pmf is shown for p = 0.1 and n = 5 and, similar to this problem, for p = 0.1 and n = 25.

Because p < 0.1, use the Poisson distribution to calculate a result with acceptable agreement with the binomial result. Begin by using the given data to calculate the value of the Poisson parameter, ρ = mean number of events = E(X). Begin by writing a similar probability expression prepared for the previous calculation in b.

$E(X) = np = \rho = 10(0.08) = 0.8$, the value of the Poisson parameter.

Probability expression: P(X) =1 – [P(X = 0) + P(X = 1)] based on Axiom 2 and showing the calculation usefulness of the complement.

$$P(x) = \frac{\rho^x \exp(-\rho)}{x!}$$

$$P(x=0) = \frac{\rho^0 \exp(-\rho)}{0!} = \exp(-\rho) = \exp(-0.8) = 0.449$$

$$P(x=1) = \frac{\rho^1 \exp(-\rho)}{1!} = \rho \exp(-\rho) = 0.8 \exp(-0.8) = 0.359$$

$$P(x>1) = 1 - [P(x=0) + P(x=1)] = 1 - (0.449 + 0.359) = 0.192 \sim 0.19$$

So Pr(x > 1) agrees with the binomial result within 2 significant digits (2 s.d.).

9. Suppose a process produces electronic components, 20% of which are defective.
    a. Find the distribution of x (the number of defective components) in a sample size of five.
       Hint: Since probability of a defective component is greater than 0.1, using Poisson distribution would result in greater uncertainty than use of Binomial distribution. Simply write the expression for the Binomial distribution, identifying the values of its parameters.
    b. Given that the sample contains atleast 3 defective components ($x \geq 3$), find the probability that 4 components are defective.
       Hint: You are to find Pr($x=4 \mid x \geq 3$)

       Solution:

Binomial Distribution with n = 5, p = 0.2 (1–p) = 0.8

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

$$P((x = 4 \mid x \geq 3) = \frac{P(x \geq 3 \mid x = 4)P(x = 4)}{P(x \geq 3)}$$

$$P(x \geq 3 \mid x = 4) = 1$$

$$\Pr(x = 4 \mid x \geq 3) = \frac{\Pr(x = 4)}{\Pr(x \geq 3)}$$

$$\Pr(x = 4 \mid x \geq 3) = \frac{\left(\frac{5}{4}\right) 0.2^4\ 0.8^1}{\left(\frac{5}{3}\right) 0.2^3\ 0.8^2 + \left(\frac{5}{4}\right) 0.2^4\ 0.8^1 + \left(\frac{5}{5}\right) 0.2^5\ 0.8^0}$$

$$\Pr(x = 4 \mid x \geq 3) = \frac{0.00640}{(0.05120) + (0.00640) + (0.00032)} = \frac{0.00640}{0.05792}$$

$$\boxed{\Pr(x = 4 \mid x \geq 3) = 0.1105}$$