Stat 50, Answers to Test 1 Problems, Spring 2025

1. The following contingency table includes the number of students in an upper division course (with a total of 80 students enrolled) according to the majors of the students and whether they passed the course:

		Major	
Performance	Math	Physics	Comp. Science
Passed	18	15	21
Didn't pass	7	10	9

A student is randomly selected from this list. Answer the following questions based on the table. Make sure to show the relevant fractions explicitly and use at least three decimal places of rounding.

(a) (2 pts) Find the probability that the student either passed the course or student was not a math major.

Answer. Let A be the event that student passed the course, and M be the event that the student is a math major. Then, we want to compute $P(A \cup M^c)$. One can use the addition rule for unions or consider the complementary event approach. With the latter approach, $(A \cup M^c)^c = A^c \cap M$ (by DeMorgan's rule), and so $P(A \cup M^c) = 1 - P(A^c \cap M) = 1 - \frac{7}{80} = \frac{73}{80} = 0.9125$.

(b) (2 pts) Determine if the following events are statistically independent? Justify your answer.

A: The student was a computer science major

B: The student passed the course.

Answer. From the table, we have $P(A \cap B) = \frac{21}{80} = 0.2625$, $P(A) = \frac{30}{80} = \frac{3}{8}$ and $P(B) = \frac{54}{80} = \frac{27}{40}$. So, $P(A) \cdot P(B) = \frac{3}{8}\frac{27}{40} \simeq 0.25313$. Since $P(A \cap B) \neq P(A)P(B)$, the events A and B are **dependent**.

2. The records of a hospital that specializes on lung cancer show that 12% of the patients who were referred to the hospital had lung cancer. The hospital gives each patient a test for lung cancer. Based on their statistics, in 97% of the cases where lung cancer was present, the test result was positive. Moreover, in 95% of the cases in which it was not present, the test was negative. A patient in the hospital is selected at random.

(a) (2.5 pts) What is the probability that the patient was tested negative for lung cancer?

Answer. Let C be the event that the patient has indeed lung cancer and N be the event that the test result is negative. Then, N^c refers to the event that the test result is positive. A tree diagram may help to summarize the information given. Using the complement rule, we also get $P(C^c) = 1 - P(C) = 0.88$ and $P(N|C) = 1 - P(N^c|C) = 1 - 0.97 = 0.03$.

Applying the total probability formula,

$$\begin{aligned} P(N) &= P(N|C)P(C) + P(N|C^c)P(C^c) \\ &= (0.03)(0.12) + (0.95)(0.88) \\ &= 0.8396. \end{aligned}$$

(b) (2 pts) Given that the patient was tested negative for cancer, compute the probability that the patient had cancer.

Answer. Applying the conditional probability formula and the result in part (a), or the Bayes' rule, we compute

$$P(C|N) = \frac{P(N|C)P(C)}{P(N)} = \frac{(0.03)(0.12)}{0.8396} \simeq 0.0043.$$

3. (2 pts) In a certain process, the probability of producing a defective component is 0.05. Assume that five independent components were inspected on a specific day. Find the probability that at least one of the items inspected is defective.

   Answer. Applying the complement rule and the (special) multiplication rule for independent events, $P(\text{at least one defective item}) = 1 - P(\text{no defective item in five trials}) = 1 - (1-0.05)^5 \simeq 0.2262$.

4. Let X be the number of hours that a machine is in use in a production facility. Assume that the probability density function (pdf) of X is given by $f(x) = \frac{2}{15}x$, for $1 < x < 4$ (and f is zero otherwise).

   (a) (2 pts) Compute $P(X > 2)$.

   Answer. Using the pdf of X, $P(X > 2) = \int_2^4 \frac{2}{15}x dx = \frac{4}{5}$.

   (b) (2 pts) Determine the cumulative distribution function (cdf) of X.

   Answer. Let $F(x) = \int_{-\infty}^{x} f(t)dt$ denote the cdf of X. Then, $F(x) = 0$, for $x \leq 1$. When $1 < x < 4$, $F(x) = \int_1^x \frac{2}{15}tdt = \frac{x^2-1}{15}$. Finally, $F(x) = 1$, for $x \geq 4$.

5. A facility has several solar panels but some of the panels are expected to fail within a year due to aging and other reasons. Let Y denote the number of panels that would fail within a year. The probability mass function (pmf), $p(y)$, of Y is estimated as follows:

y :	0	1	2	3	4	5
$p(y)$:	0.15	0.2	0.35	0.2	0.07	0.03

   (a) (1.5 pts) What is the probability that at least 2 panels would fail within a year?
   Answer. Using the probability table above, $P(X \geq 2) = p(2) + p(3) + p(4) + p(5) = 0.65$. Complement rule may also be considered: $1 - P(X < 2) = 1 - [p(0) + p(1)] = 1 - 0.35 = 0.65$.

   (b) (2 pts) Compute the mean value of Y. Include the units.

   Answer. Applying the discrete r.v. mean formula to this distribution, $\mu = \sum_{x=0}^{5} xp(x) = 0 + 0.2 + 0.7 + 0.6 + 0.28 + 0.15 = 1.93$ panels/year, on average.

   (c) (2 pts) If the variance of Y is $\sigma_Y^2 = 1.5451$, then compute the z-score of $y = 5$ and briefly explain its meaning in context.
   Answer. First obtain the standard deviation: $\sigma_Y = \sqrt{1.5451} = 1.243$. Then, the z-score is $z = \frac{5-1.93}{1.243} \simeq 2.47$, indicating it is about 2.47 standard deviations larger than the avrage number of failures. So, if 5 panels fail within a year, it is somewhat unusual since it is more than two standard deviations from the mean.

BONUS. (1 pt) Let Y_1 and Y_2 have the same distribution as the r.v. Y as above. Determine the mean of the r.v.$W = \frac{Y_1+Y_2}{2}$.
Answer. Note that W represents the sample mean as a random variable. Applying the properties of the mean (or expected value) for linear combinations, we get $E[W] = \frac{1}{2}E[Y_1+Y_2] = \frac{1}{2}(E[Y_1] + E[Y_2]) = \frac{1}{2}(1.93 + 1.93) = 1.93$ panels.