STAT 400
UIUC

Practice Problems #3

Stepanov
Dalpiaz

SOLUTIONS

The following are a number of practice problems that may be *helpful* for completing the homework, and will likely be **very useful** for studying for exams.

1. **1.2-10** **1.3-14** 1.3-14

Prove (show) that

$$\binom{n-1}{r} + \binom{n-1}{r-1} = \binom{n}{r}.$$

(Pascal's equation).

$$\binom{n-1}{r} + \binom{n-1}{r-1} = \frac{(n-1)!}{r! \cdot (n-r-1)!} + \frac{(n-1)!}{(r-1)! \cdot (n-r)!}$$

$$= \frac{(n-1)!}{(r-1)! \cdot (n-r-1)!} \cdot \left[\frac{1}{r} + \frac{1}{n-r} \right]$$

$$= \frac{(n-1)!}{(r-1)! \cdot (n-r-1)!} \cdot \frac{n}{r \cdot (n-r)} = \frac{n!}{r! \cdot (n-r)!} = \binom{n}{r}.$$

2. **1.2-16** **1.3-20** 1.3-20

A box of candy hearts contains 52 hearts, of which 19 are white, 10 are tan, 7 are pink, 3 are purple, 5 are yellow, 2 are orange, and 6 are green. If you select 9 pieces of candy randomly from the box, without replacement, give the probability that

a) Three of the hearts are white.

$$\frac{\binom{19}{3}\cdot\binom{52-19}{6}}{\binom{52}{9}} = \frac{969\cdot 1{,}107{,}568}{3{,}679{,}075{,}400} = \frac{102{,}486}{351{,}325} \approx 0.291713.$$

b) Three are white, 2 are tan, 1 is pink, 1 is yellow, and 2 are green.

$$\frac{\binom{19}{3}\cdot\binom{10}{2}\cdot\binom{7}{1}\cdot\binom{3}{0}\cdot\binom{5}{1}\cdot\binom{2}{0}\cdot\binom{6}{2}}{\binom{52}{9}} = \frac{969\cdot 45\cdot 7\cdot 1\cdot 5\cdot 1\cdot 15}{3{,}679{,}075{,}400} = \frac{7{,}695}{1{,}236{,}664} \approx 0.0062224.$$

3. Peter takes Computer Science classes, though not to learn, but to meet smart girls. There are 15 other students in the class with Peter, 7 of them are girls. During the semester, students will be working on a project in teams of 4 students. Suppose the students are divided into teams at random.

Peter's team = Peter + 3 (randomly chosen) students.

Students are selected at random without replacement.

a) Find the probability that at least 2 out of 3 students on Peter's team are girls.

$$\frac{{}_7C_2 \cdot {}_8C_1}{{}_{15}C_3} + \frac{{}_7C_3 \cdot {}_8C_0}{{}_{15}C_3} = \frac{21 \cdot 8}{455} + \frac{35 \cdot 1}{455} = \frac{203}{455} = \mathbf{\frac{29}{65}} \approx 0.446154.$$

b) Find the probability that there is at least 1 girl on Peter's team.

$$1 - \frac{{}_7C_0 \cdot {}_8C_3}{{}_{15}C_3} = 1 - \frac{1 \cdot 56}{455} = \frac{399}{455} = \mathbf{\frac{57}{65}} \approx 0.876923.$$

OR

$$\frac{{}_7C_1 \cdot {}_8C_2}{{}_{15}C_3} + \frac{{}_7C_2 \cdot {}_8C_1}{{}_{15}C_3} + \frac{{}_7C_3 \cdot {}_8C_0}{{}_{15}C_3} = \frac{7 \cdot 28}{455} + \frac{21 \cdot 8}{455} + \frac{35 \cdot 1}{455}$$

$$= \frac{399}{455} = \mathbf{\frac{57}{65}} \approx 0.876923.$$

c) Find the probability that at most 2 out of 3 students on Peter's team are girls.

$$1 - \frac{{}_7C_3 \cdot {}_8C_0}{{}_{15}C_3} = 1 - \frac{35 \cdot 1}{455} = \frac{420}{455} = \mathbf{\frac{12}{13}} \approx 0.923077.$$

OR

$$\frac{{}_7C_0 \cdot {}_8C_3}{{}_{15}C_3} + \frac{{}_7C_1 \cdot {}_8C_2}{{}_{15}C_3} + \frac{{}_7C_2 \cdot {}_8C_1}{{}_{15}C_3} = \frac{1 \cdot 56}{455} + \frac{7 \cdot 28}{455} + \frac{21 \cdot 8}{455}$$

$$= \frac{420}{455} = \mathbf{\frac{12}{13}} \approx 0.923077.$$

4. A small grocery store had 10 cartons of milk, 2 of which were sour.

a) If David is going to buy the sixth carton of milk sold that day at random, compute the probability that he selects a carton of sour milk.

$$P\begin{pmatrix}\text{first 5:}\\ \text{5 fresh, 0 sour}\end{pmatrix}\times\frac{2}{5} + P\begin{pmatrix}\text{first 5:}\\ \text{4 fresh, 1 sour}\end{pmatrix}\times\frac{1}{5} + P\begin{pmatrix}\text{first 5:}\\ \text{3 fresh, 2 sour}\end{pmatrix}\times\frac{0}{5}$$

$$= \frac{\binom{8}{5}\cdot\binom{2}{0}}{\binom{10}{5}}\times\frac{2}{5} + \frac{\binom{8}{4}\cdot\binom{2}{1}}{\binom{10}{5}}\times\frac{1}{5} = \frac{56\cdot 1}{252}\times\frac{2}{5} + \frac{70\cdot 2}{252}\times\frac{1}{5} = \mathbf{\frac{1}{5}}.$$

b) If six cartons of milk are sold that day at random, what is the probability that exactly one carton of sour milk is sold.

$$\frac{\binom{8}{5}\cdot\binom{2}{1}}{\binom{10}{6}} = \frac{56\cdot 2}{210} = \mathbf{\frac{8}{15}}.$$

5. Suppose the number of boxes of Hammermill® paper used by Anytown College Statistics & Probability Department each month is random and has the following probability distribution:

x	$f(x)$	$x \cdot f(x)$	$x^2 \cdot f(x)$
0	0.1	0	0
1	0.1	0.1	0.1
2	0.3	0.6	1.2
3	0.3	0.9	2.7
4	0.2	0.8	3.2
		2.4	7.2

Suppose at the end of each month the department orders the same number of boxes as was used during the month. Suppose each box costs $25. The department has to pay a $5 delivery fee (the delivery fee does not depend on the number of boxes ordered). Then the monthly "paper" bill is $Y = 25 \cdot X + 5$. Find Anytown College Statistics & Probability Department's average monthly "paper" bill and its standard deviation.

$$E(X) = \sum_{\text{all } x} x \cdot f(x) = \mathbf{2.4}.$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \sum_{\text{all } x} x^2 \cdot f(x) - [E(X)]^2 = 7.2 - (2.4)^2 = \mathbf{1.44}.$$

$$SD(X) = \sqrt{1.44} = \mathbf{1.2}.$$

$$E(Y) = 25 \cdot E(X) + 5 = 25 \cdot 2.4 + 5 = \mathbf{\$65}.$$

$$SD(Y) = |25| \cdot SD(X) = 25 \cdot 1.2 = \mathbf{\$30}.$$

6. Suppose we roll a pair of fair 6-sided dice. Let X denote the maximum (the largest) of the outcomes on the two dice. Construct the probability distribution of X and compute its expected value.

(1 , 1)	(1 , 2)	(1 , 3)	(1 , 4)	(1 , 5)	(1 , 6)
(2 , 1)	(2 , 2)	(2 , 3)	(2 , 4)	(2 , 5)	(2 , 6)
(3 , 1)	(3 , 2)	(3 , 3)	(3 , 4)	(3 , 5)	(3 , 6)
(4 , 1)	(4 , 2)	(4 , 3)	(4 , 4)	(4 , 5)	(4 , 6)
(5 , 1)	(5 , 2)	(5 , 3)	(5 , 4)	(5 , 5)	(5 , 6)
(6 , 1)	(6 , 2)	(6 , 3)	(6 , 4)	(6 , 5)	(6 , 6)

x	$f(x)$
1	$1/36$
2	$3/36$
3	$5/36$
4	$7/36$
5	$9/36$
6	$11/36$

$x \times f(x)$
$1/36$
$6/36$
$15/36$
$28/36$
$45/36$
$66/36$

$$E(X) = 161/36 \approx 4.472222.$$

7. Consider $f(x) = c(x+1)^2, \quad x = 0, 1, 2, 3.$

a) Find c such that $f(x)$ satisfies the conditions of being a p.m.f. for a random variable X.

$$1 = \sum_{\text{all } x} f(x) = f(0) + f(1) + f(2) + f(3) = c + 4c + 9c + 16c = 30c.$$

$$\Rightarrow \quad c = \mathbf{\frac{1}{30}}.$$

b) Find the expected value of X.

$$\mathrm{E}(\mathrm{X}) = \sum_{\text{all } x} x \cdot f(x) = 0 \times \frac{1}{30} + 1 \times \frac{4}{30} + 2 \times \frac{9}{30} + 3 \times \frac{16}{30} = \mathbf{\frac{7}{3}} \approx 2.33333.$$

c) Find the standard deviation of X.

$$\mathrm{E}(\mathrm{X}^2) = \sum_{\text{all } x} x^2 \cdot f(x) = 0 \times \frac{1}{30} + 1 \times \frac{4}{30} + 4 \times \frac{9}{30} + 9 \times \frac{16}{30} = \frac{92}{15} \approx 6.13333.$$

$$\mathrm{Var}(\mathrm{X}) = \mathrm{E}(\mathrm{X}^2) - [\mathrm{E}(\mathrm{X})]^2 = \frac{92}{15} - \frac{49}{9} = \frac{31}{45} \approx 0.68889.$$

$$\mathrm{SD}(\mathrm{X}) = \sqrt{\mathrm{Var}(\mathrm{X})} = \sqrt{\frac{31}{45}} \approx \mathbf{0.83}.$$

8.

a) Let X be a discrete random variable with p.m.f.

$$f(k) = \frac{c}{a^k}, \quad k = 2, 3, 4, 5, 6, \dots, \qquad \text{where } c = a(a-1).$$

Recall (Homework #1 Problem 7): this a valid probability distribution.

Find $\mu_X = \text{E}(\text{X})$.

$$\text{E}(\text{X}) = \sum_{\text{all } x} x \cdot f(x) = \sum_{k=2}^{\infty} k \cdot \frac{a(a-1)}{a^k} = \sum_{k=2}^{\infty} k \cdot \frac{a-1}{a^{k-1}}.$$

$$\text{E}(\text{X}) = 2 \cdot \frac{a-1}{a^1} + 3 \cdot \frac{a-1}{a^2} + 4 \cdot \frac{a-1}{a^3} + 5 \cdot \frac{a-1}{a^4} + 6 \cdot \frac{a-1}{a^5} + \dots$$

$$\frac{1}{a} \cdot \text{E}(\text{X}) = \qquad 2 \cdot \frac{a-1}{a^2} + 3 \cdot \frac{a-1}{a^3} + 4 \cdot \frac{a-1}{a^4} + 5 \cdot \frac{a-1}{a^5} + \dots$$

$$\Rightarrow \quad \left(1 - \frac{1}{a}\right) \cdot \text{E}(\text{X}) = 2 \cdot \frac{a-1}{a^1} + \frac{a-1}{a^2} + \frac{a-1}{a^3} + \frac{a-1}{a^4} + \frac{a-1}{a^5} + \dots$$

$$= \frac{a-1}{a} + \frac{a(a-1)}{a^2} + \frac{a(a-1)}{a^3} + \frac{a(a-1)}{a^4} + \frac{a(a-1)}{a^5} + \frac{a(a-1)}{a^6} + \dots.$$

$$= \frac{a-1}{a} + 1 = \frac{2a-1}{a}.$$

$$\frac{a-1}{a} \cdot \text{E}(\text{X}) = \frac{2a-1}{a}.$$

Therefore, $\text{E}(\text{X}) = \dfrac{2a-1}{a-1} = 1 + \dfrac{a}{a-1} = 2 + \dfrac{1}{a-1}$.

OR

$$\text{E}(\text{X}) = \sum_{\text{all } x} x \cdot f(x) = \sum_{k=2}^{\infty} k \cdot \frac{a(a-1)}{a^k} = \sum_{k=1}^{\infty} (k+1) \cdot \frac{a(a-1)}{a^{k+1}}$$

$$= \sum_{k=1}^{\infty} (k+1) \cdot \frac{1}{a^{k-1}} \cdot \frac{a-1}{a} = \sum_{k=1}^{\infty} (k+1) \cdot \left(\frac{1}{a}\right)^{k-1} \cdot \left(1 - \frac{1}{a}\right) = \mathrm{E}(\mathrm{Y}+1),$$

where Y has a Geometric distribution with probability of "success" $p = 1 - \frac{1}{a} = \frac{a-1}{a}$.

$$\Rightarrow \quad \mathrm{E}(\mathrm{X}) = \mathrm{E}(\mathrm{Y}) + 1 = \frac{a}{a-1} + 1 = \frac{2a-1}{a-1} = 2 + \frac{1}{a-1}.$$

b) Let X be a discrete random variable with p.m.f.

$$f(k) = c\,\frac{2^k}{k!}, \quad k = 2, 3, 4, 5, 6, \ldots, \qquad \text{where } c = \frac{1}{e^2 - 3}.$$

Recall (Homework #1 Problem 8): this a valid probability distribution.

Find $\mu_{\mathrm{X}} = \mathrm{E}(\mathrm{X})$.

$$\mathrm{E}(\mathrm{X}) = \sum_{\text{all } x} x \cdot f(x) = \sum_{k=2}^{\infty} k \cdot \frac{1}{e^2 - 3} \frac{2^k}{k!} = \frac{1}{e^2 - 3} \sum_{k=2}^{\infty} \frac{2^k}{(k-1)!}$$

$$= \frac{1}{e^2 - 3} \sum_{n=1}^{\infty} \frac{2^{n+1}}{n!} = \frac{2}{e^2 - 3} \sum_{n=1}^{\infty} \frac{2^n}{n!} = \frac{2\left(e^2 - 1\right)}{e^2 - 3} \approx 2.91136.$$

9 – 10. An oil company believes that the probability of existence of an oil deposit in a certain drilling area is 0.30. Suppose it would cost \$100,000 to drill a well. If an oil deposit does exist, the company's profit will be \$700,000 (the drilling costs not included). A seismic test that would cost \$20,000 is being considered to clarify the likelihood of the presence of oil. The proposed seismic test has the following reliability: when oil does exist in the testing area, the test will indicate so 90% of the time; when oil does not exist in the test area, 20% of the time the test will erroneously indicate that it does exist. There are four possible "states of nature":

θ_1 = an oil deposit does exist and the test result is positive,

θ_2 = an oil deposit does exist, but (and) the test result is negative,

θ_3 = an oil deposit does not exist, but (and) the test result is positive,

θ_4 = an oil deposit does not exist and the test result is negative.

The company can take two possible actions:

a_1 = drill a well without performing the test,

a_2 = perform the test and drill a well only if the test shows presence of oil.

9. a) Find the probabilities of all four states of nature.
That is, find $P(\theta_1)$, $P(\theta_2)$, $P(\theta_3)$, and $P(\theta_4)$.

$P(\text{Oil}) = 0.30$, $P(+ \mid \text{Oil}) = 0.90$, $P(+ \mid \text{No Oil}) = 0.20$.

$P(\theta_1) = P(\text{Oil} \cap +) = P(\text{Oil}) \times P(+ \mid \text{Oil}) = 0.30 \times 0.90 = \mathbf{0.27}$.

$P(\theta_2) = P(\text{Oil} \cap -) = P(\text{Oil}) \times P(- \mid \text{Oil}) = 0.30 \times 0.10 = \mathbf{0.03}$.

$P(\theta_3) = P(\text{No Oil} \cap +) = P(\text{No Oil}) \times P(+ \mid \text{No Oil}) = 0.70 \times 0.20 = \mathbf{0.14}$.

$P(\theta_4) = P(\text{No Oil} \cap -) = P(\text{No Oil}) \times P(- \mid \text{No Oil}) = 0.70 \times 0.80 = \mathbf{0.56}$.

b) Suppose the test shows presence of oil. What is the probability that an oil deposit does exist?

$$P(\text{Oil} \mid +) = \frac{P(\text{Oil} \cap +)}{P(+)} = \frac{0.27}{0.27 + 0.14} = \frac{0.27}{0.41} \approx \mathbf{0.6585366}.$$

10. c) Construct the payoff table (profit table) for this problem. That is, find the company's profit for each possible action and each possible state of nature.

	θ_1 Oil +	θ_2 Oil –	θ_3 No Oil +	θ_4 No Oil –
a_1 drill w/o test	– 100,000 700,000 **600,000**	– 100,000 700,000 **600,000**	– 100,000 **– 100,000**	– 100,000 **– 100,000**
a_2 drill only if +	– 20,000 – 100,000 700,000 **580,000**	– 20,000 **– 20,000**	– 20,000 – 100,000 **– 120,000**	– 20,000 **– 20,000**

d) Find the expected payoff (expected profit, EP) for both actions and determine the optimal action.

EP(a_1) = 600,000 × 0.27 + 600,000 × 0.03 + (–100,000) × 0.14 + (–100,000) × 0.56
= **$110,000**.

EP(a_2) = 580,000 × 0.27 + (–20,000) × 0.03 + (–120,000) × 0.14 + (–20,000) × 0.56
= **$128,000**.

Optimal action = $\mathbf{a_2}$ (perform the test and drill a well only if the test shows presence of oil) is the optimal action, it has a higher expected payoff.

For fun:

Suppose the probability of existence of an oil deposit in a certain drilling area is unknown, p.

$P(\theta_1) = P(\text{Oil} \cap +) = P(\text{Oil}) \times P(+ \mid \text{Oil}) = p \times 0.90.$

$P(\theta_2) = P(\text{Oil} \cap -) = P(\text{Oil}) \times P(- \mid \text{Oil}) = p \times 0.10.$

$P(\theta_3) = P(\text{No Oil} \cap +) = P(\text{No Oil}) \times P(+ \mid \text{No Oil}) = (1 - p) \times 0.20.$

$P(\theta_4) = P(\text{No Oil} \cap -) = P(\text{No Oil}) \times P(- \mid \text{No Oil}) = (1 - p) \times 0.80.$

	θ_1 Oil +	θ_2 Oil –	θ_3 No Oil +	θ_4 No Oil –
a_1 drill w/o test	– 100,000 700,000 **600,000**	– 100,000 700,000 **600,000**	– 100,000 **– 100,000**	– 100,000 **– 100,000**
a_2 drill only if +	– 20,000 – 100,000 700,000 **580,000**	– 20,000 **– 20,000**	– 20,000 – 100,000 **– 120,000**	– 20,000 **– 20,000**
a_3 do nothing	**0**	**0**	**0**	**0**

$EP(a_1) = 600{,}000 \times p \times 0.90 + 600{,}000 \times p \times 0.10$

$+ (-100{,}000) \times (1 - p) \times 0.20 + (-100{,}000) \times (1 - p) \times 0.80$

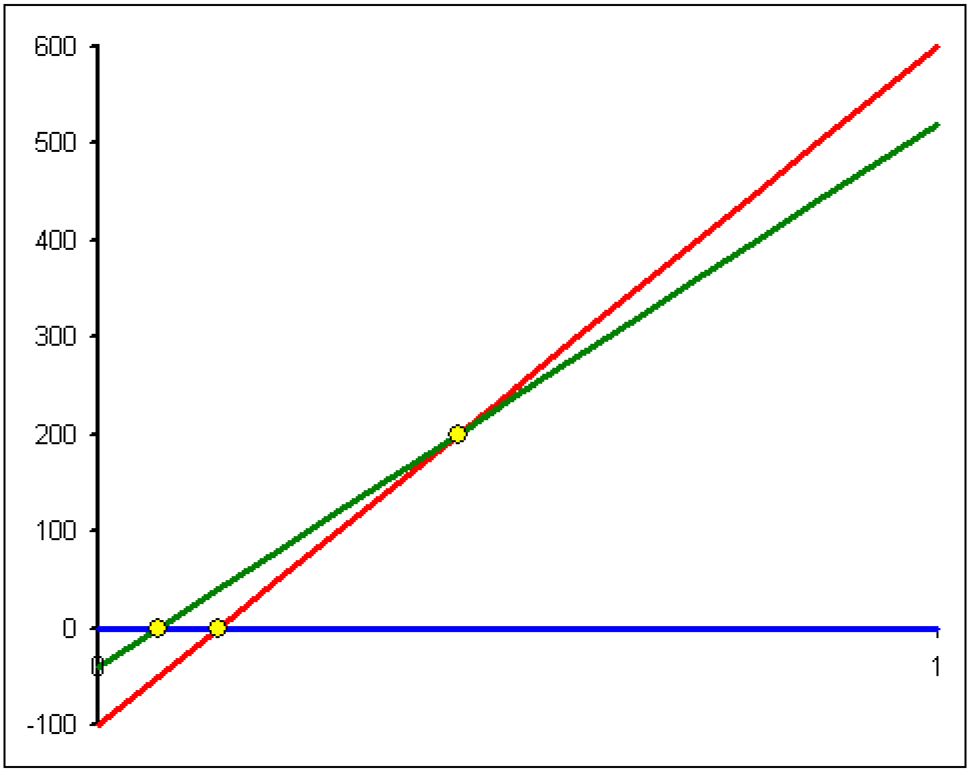
$= 700{,}000 \times p - 100{,}000.$

$EP(a_2) = 580{,}000 \times p \times 0.90 + (-20{,}000) \times p \times 0.10$

$+ (-120{,}000) \times (1 - p) \times 0.20 + (-20{,}000) \times (1 - p) \times 0.80$

$= 560{,}000 \times p - 40{,}000.$

$EP(a_3) = 0.$



$EP(a_1) = EP(a_2)$.

$700{,}000\,p - 100{,}000 = 560{,}000\,p - 40{,}000$.

$140{,}000\,p = 60{,}000$. $\Rightarrow$ $p = 3/7$.

$EP(a_1) = EP(a_3)$.

$700{,}000\,p - 100{,}000 = 0$.

$\Rightarrow$ $p = 1/7$.

$EP(a_2) = EP(a_3)$.

$560{,}000\,p - 40{,}000 = 0$.

$\Rightarrow$ $p = 1/14$.

If $p < 1/14$, $\mathbf{a_3}$ is optimal.

If $1/14 < p < 3/7$, $\mathbf{a_2}$ is optimal.

If $p > 3/7$, $\mathbf{a_1}$ is optimal.