Math 461 B, Spring 2009
Final Exam Solutions and Comments

1. An instructor wants to create an exam consisting of 5 problems and covering 6 sections of the text. To this end, he first makes up 10 problems for each of the 6 sections, and then selects at random 5 *different* problems from these 60 problems.

   (a) What is the probability that the problems on the exam are all from different sections (i.e., that no section has more than one problem on the exam)?

   10 pts

   Solution. [This problem is just like as the final exam scheduling problem discussed in class (involving 6 final exam days with 3 slots per day)]
   The total number of ways of selecting 5 (different) problems out of 60, taking order into account, is $60 \cdot 59 \cdot 58 \cdot 57 \cdot 56$. The number of such selections in which each problem is from a different section is $60 \cdot 50 \cdot 40 \cdot 30 \cdot 20$ since there are 60 choices for the first problem, 50 for the second problem (since it can't be one of the 10 problems from the section of problem 1), 40 for the third problem, etc. Thus, the probability that all problems are from a different sections is

$$\frac{60 \cdot 50 \cdot 40 \cdot 30 \cdot 20}{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56} = \frac{\binom{6}{5}10^5}{\binom{60}{5}}$$

   [An alternative approach is to count unordered samples. Then the counts are $\#(S) = \binom{60}{5}$, and $\#(A) = \binom{6}{5} \cdot 10^5$ (pick 5 sections out of 6, then pick from each of these 5 sections 1 out of 10 problems). The resulting probability, $\binom{6}{5}10^5/\binom{60}{5}$, is the same as the one obtained above.]

   (b) What is the probability that the problems on the exam are all from the same section?

   10 pts

   Solution. The total number of ways of selecting 5 (different) problems out of 60, taking order into account, is again $60 \cdot 59 \cdot 58 \cdot 57 \cdot 56$. The number of such selections in which all problems are from the same section is $6 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6$ (6 ways to pick a section, and $10 \cdot 9 \cdot 8 \cdot 7 \cdot 6$ ways to pick an *ordered* sample of 5 *without replacement* out of the 10 problems. Hence, the probability is

$$\frac{6 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56} = \frac{6\binom{10}{5}}{\binom{60}{5}}$$

   (c) What is the expected number of sections from which there is a problem on the exam?

   10 pts

   Solution. We need to compute $E(X)$, where X is the number of sections that have a problem on the exam. The only feasible way to do this is by the indicator method, since the individual probabilities $P(X = x)$ are very difficult to compute. We define the events

$$A_i = \text{“Section } i \text{ has a problem on the exam”} \qquad (i = 1, 2, \ldots, 6).$$

   Then X is equal to the number of A_i's that occur, so by the indicator method we have

$$E(X) = \sum_{i=1}^{6} P(A_i).$$

   To compute $P(A_i)$, we use the complement trick:

$$\begin{aligned} P(A_i) &= 1 - P(A_i^c) = 1 - P(\text{no problem from section } i \text{ is on the exam}) \\ &= 1 - P(\text{5 chosen problems are among the 50 (out of 24) not in section } i) \\ &= 1 - \frac{50 \cdot 49 \cdots 46}{60 \cdot 59 \cdots 56} \end{aligned}$$

   so

$$E(X) = 6\left(1 - \frac{50 \cdot 49 \cdots 46}{60 \cdot 59 \cdots 56}\right).$$

2. Suppose $P(A) = 1/4$, $P(B) = 1/3$, and $A \subset B$.

   (a) Find $P(A \mid B)$.

   6 pts **Solution.** Since $A \subset B$, we have $AB = A$, so

$$P(A \mid B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)} = \frac{1/4}{1/3} = \boxed{\frac{3}{4}}$$

   (b) Find $P(B \mid A)$.

   6 pts **Solution.**

$$P(B \mid A) = \frac{P(BA)}{P(A)} = \frac{P(A)}{P(A)} = \boxed{1}$$

   (c) Find $P(B \mid A^c)$

   6 pts **Solution.**

$$P(B \mid A^c) = \frac{P(BA^c)}{P(A^c)} = \frac{P(B) - P(A)}{1 - P(A)} = \frac{1/3 - 1/4}{1 - 1/4} = \boxed{\frac{1}{9}}$$

   (d) Find $P(B^c \mid A^c)$

   6 pts **Solution.**

$$P(B^c \mid A^c) = 1 - P(B \mid A^c) = 1 - \frac{1}{9} = \boxed{\frac{8}{9}}$$

   (e) Are A and B^c independent? Justify your answer.

   6 pts **Solution.** Since $A \subset B$, we have $A \cap B^c = \emptyset$, so $P(AB^c) = 0$. On the other hand, $P(A)P(B^c) = (1/4)(2/3) \neq 0$, so $P(A)P(B^c) \neq P(AB^c)$. Hence A and B^c are not independent.

3. A die is rolled repeatedly. Let X denote the number of the roll at which the **third** six occurs.

   (a) Find $P(X > 461)$ **without using the result of part (b)** (i.e., without using any formulas for $P(X = k)$). Your answer can be in "raw" form, but should be such that a numerical value could be easily computed with a basic calculator. A sum involving a large number of terms would not qualify.

   10 pts **Solution.** The event $X > 461$ means, by the definition of X, that the third six occurs after roll 461. But this is equivalent to saying that there are at most 2 sixes in the first 461 rolls. The probability for this event is a standard success/failure probability, namely

$$P(\leq 2 \text{ sixes in 461 rolls}) = \left(\frac{5}{6}\right)^{461} + \binom{461}{1}\left(\frac{5}{6}\right)^{460}\left(\frac{1}{6}\right) + \binom{461}{2}\left(\frac{5}{6}\right)^{359}\left(\frac{1}{6}\right)^{2}.$$

   Comment: The key to this problem is the rephrasing of the event "$X > 461$" as "at most two sixes in the first 461 rolls". This type of reasoning has come up before (e.g., in Problem 5 of HW 3, or in connection with the birthday problem). Of course, the probability $P(X > 461)$ can be written as a sum of the 461 probabilities $P(X = n)$ for $n = 1, 2, \ldots, 461$, but this is not a practical way to solve the problem, and it is an acceptable solution.

(b) Find the p.m.f. of X. Be sure to specify the range.

10 pts

Solution.

Values of X (range): $3, 4, \dots$.

Computation of probabilities: For $n = 3, 4, \dots$, we have

$$\begin{aligned} P(X = n) &= P(\text{third six occurs at trial } n) \\ &= P(\text{exactly two sixes in trials } 1, 2, \dots, n-1 \text{ and one six in trial } n) \\ &= \binom{n-1}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{n-3}\left(\frac{1}{6}\right) = \binom{n-1}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^{n-3} \end{aligned}$$

(Alternatively, this could have been derived from the negative binomial distribution.)

Comment: An answer like $P(X = n) = \binom{n}{3}(1/6)^3(5/6)^{n-3}$ would be completely off target, as this is a straight S/F probability, namely the probability for getting three sixes in n rolls, which does not take into account the requirement that the third six has to occur at the n-th roll. (This fallacy come up in several homework problems and was pointed out repeatedly in that context.)

(c) Find the expectation of X. (As always, you have to explain/justify your answer, e.g., by stating precisely which formula/theorem/property you are using!)

10 pts

Solution. In S/F language (with S denoting a six and F denoting a non-six), X is the trial at which the 3rd success occurs. But such a random variable has negative binomial distribution with parameter $r = 3$ and $p = 1/6$. *By the formula for the expectation of a negative binomial distribution,* $E(X) = 3/p = 3/(1/6) = \boxed{18}$. (The required justification is the phrase in italics.)

4. Let X be a random variable with density (p.d.f.)

$$f(x) = (1/2)e^{-x/2}, \quad 0 < x < \infty.$$

(a) Find $M_X(t)$, the moment-generating function (mgf) of X, and determine the range of values t for which the mgf exists. (You must carry out the necessary calculations. No credit for citing a memorized formula for the mgf.)

10 pts

Solution. [This is essentially Problem 8 in HW 10.]

$$\begin{aligned} M(t) &= \int_0^\infty e^{tx}(1/2)e^{-x/2}dx \\ &= \frac{1}{2}\int_0^\infty \exp\left\{-x\left(\frac{1}{2} - t\right)\right\} dx \\ &= \frac{1}{2(\frac{1}{2} - t} = \boxed{\frac{1}{1 - 2t}} \end{aligned}$$

provided $\boxed{t < 1/2}$. If $t \geq 1/2$, the above integral diverges, so $M_X(t)$ does not exist in this range.

10 pts (b) Find $P(X \le 6 \mid X \ge 3)$.

Solution. We compute:

$$\begin{aligned} P(X \ge 3) &= \int_3^\infty (1/2)e^{-x/2}dx = e^{-3/2}, \\ P(3 \le X \le 6) &= \int_3^6 (1/2)e^{-x/2}dx = e^{-3/2} - e^{-6/2}, \\ P(X \le 6 \mid X \ge 3) &= \frac{P(3 \le X \le 6)}{P(X \ge 3)} \\ &= \frac{e^{-3/2} - e^{-6/2}}{e^{-3/2}} = \boxed{1 - e^{-3/2}} \end{aligned}$$

10 pts (c) Let $Y = \sqrt{X}$. Find the p.d.f. of Y.

Solution. We use the change of variables technique, computing first the c.d.f.'s of X and Y, then differentiating the latter to get the p.d.f. of Y.

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(\sqrt{X} \le y) = P(X \le y^2) = F_X(y^2), \\ f_Y(y) &= \frac{d}{dy}F_X(y^2) = F_X'(y^2)2y = f_X(y^2)2y \\ &= \boxed{ye^{-y^2/2}, \quad 0 < y < \infty} \end{aligned}$$

5. Suppose X and Y are discrete random variables with values $1, 2, 3$ each and joint p.m.f. given by

$$f(x,y) = \begin{cases} 1/9 & \text{if } x = y \\ 2/9 & \text{if } x < y \\ 0 & \text{if } x > y \end{cases}$$

for $x, y = 1, 2, 3$.

10 pts (a) Find the marginal p.m.f.'s of X and Y.

Solution. The matrix representation of the joint distribution is as follows, with marginal distributions of X and Y given in the last column and last row.

$X \setminus Y$	1	2	3	$p_X(x)$
1	1/9	2/9	2/9	5/9
2	0	1/9	2/9	3/9
3	0	0	1/9	1/9
$p_Y(y)$	1/9	3/9	5/9	

10 pts (b) Find $E(XY)$.

Solution.

$$E(XY) = \sum xyP(x,y) = \frac{1}{9}(1\cdot 1 + 2\cdot 2 + 3\cdot 3) + \frac{2}{9}(1\cdot 2 + 1\cdot 3 + 2\cdot 3) = \boxed{4}$$

(c) Find the conditional p.m.f. of Y given $X = 1$, and represent it in the form of a distribution table (i.e., a 2-row table with the first row listing the values and the second row the associated probabilities).

10 pts

Solution. The conditional p.m.f. of Y given $X = 1$ is given by

$$p(Y = y \mid X = 1) = \frac{p(1, y)}{p_X(1)}.$$

Using the above values for $p(x, y)$ and $p_X(x)$, we get

y	1	2	3
$P(Y = y \mid X = 1)$	1/5	2/5	2/5

6. Suppose that X is uniformly distributed on the interval $[0, 1]$ and that, given $X = x$, Y is uniformly distributed on the interval $[1 - x, 1]$.

(a) Determine the joint density $f(x, y)$. (Be sure to specify the range.)

10 pts

Solution. Since X is uniformly distributed on $[0, 1]$, we have $f_X(x) = 1,\ 0 \le x \le 1$. Similarly, since, given $X = x$, Y is uniformly distributed on $[1 - x, 1]$, the conditional density of Y given $X = x$ is $1/(1 - (1 - x)) = 1/x$ on the interval $[1 - x, 1]$; i.e., $f_{Y|X}(y|x) = 1/x,\ 1 - x \le y \le 1$ for $0 \le x \le 1$. Thus

$$f(x, y) = f_X(x) f_{Y|X}(y|x) = \boxed{\frac{1}{x}, \quad 0 < x < 1, 1 - x < y < 1}$$

(b) Find the probability $P(X \ge 1/2, Y \ge 1/2)$. (As usual, you can leave your answer in raw form, such as $1 - 1/e$.)

10 pts

Solution. This can be computed either as an integral of the joint density over appropriate region, or as a single integral over the marginal density $f_Y(y)$ from $1/2$ to 1. Since we know the joint density from part (a), the first method is the more natural one. We get (see sketch for the integration limits)

$$\begin{aligned} P(X \ge 1/2, Y \ge 1/2) &= \int_{x=1/2}^{1} \int_{y=1/2}^{1} \frac{1}{x} dy dx \\ &= \int_{x=1/2}^{1} \frac{1/2}{x} dx \\ &= \boxed{\frac{\ln 2}{2}} \end{aligned}$$

(c) Find the conditional density, $f_{X|Y}(x|1/3)$, of X given $Y = 1/3$. Be sure to specify the range.

10 pts

Solution. We first compute the marginal density $f_Y(y)$:

$$f_Y(y) = \int_{x=1-y}^{1} \frac{1}{x} dx = \ln 1 - \ln(1 - y) = -\ln(1 - y), \quad 0 \le y \le 1.$$

Hence

$$f_{X|Y}(x|1/3) = \frac{f(x, 1/3)}{f_Y(1/3)} = -\frac{1}{x \ln(1 - 1/3)} = \boxed{\frac{1}{x \ln(3/2)}, \quad 2/3 \le x \le 1}$$

(Note that, since Y is fixed at the value $Y = 1/3$, the answer should not involve the variable y, in either the formula or the range.)

7. Suppose X_1 and X_2 are independent random variables, each having density

$$f(x) = \begin{cases} 3x^{-4} & \text{for } 1 < x < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $\text{Var}(X_1)$, $\text{Var}(X_2)$, and $\text{Cov}(X_1, X_2)$. (Hint: This requires only a minimal amount of calculations.)

10 pts

Solution.

$$E(X_1) = \int_1^\infty x \cdot 3x^{-4} dx = \int_1^\infty 3x^{-3} dx = \frac{3}{2},$$
$$E(X_1^2) = \int_1^\infty x^2 \cdot 3x^{-4} dx = \int_1^\infty 3x^{-2} dx = 3,$$
$$\text{Var}(X_1) = E(X_1^2) - E(X_1)^2 = 3 - \left(\frac{3}{2}\right)^2 = \boxed{\frac{3}{4}}$$

Since X_2 has the same distribution as X_1, we have $\boxed{\text{Var}(X_2) = \text{Var}(X_1) = 3/4}$.

Since X_1 and X_2 are independent, we have $\boxed{\text{Cov}(X_1, X_2) = 0}$.

(b) Let $Z = \max(X_1, X_2)$ denote the larger (maximum) of the two random variables. Find the p.d.f. (density) of Z.

10 pts

Solution. To find the p.d.f. $f_Z(z)$ of Z, we first compute the c.d.f. $F_Z(z)$, using the "maximum trick" that came up in class on several occasions: For $z \geq 1$,

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(\max(X_1, X_2) \leq z) \\ &= P(X_1 \leq z, X_2 \leq z) = P(X_1 \leq z)P(X_2 \leq z) = F(z)^2. \end{aligned}$$

Now, $f(z) = 3z^{-4}$ for $z \geq 1$, so $F(z) = \int_1^z 3t^{-4} dt = 1 - z^{-3}$, and therefore

$$\begin{aligned} F_Z(z) &= (1 - z^{-3})^2, \\ f_Z(z) &= F_Z'(z) = 2(1 - z^{-3})(-(-3)z^{-4}) \\ &= \boxed{6z^{-4} - 6z^{-7} \quad (z \geq 1)}. \end{aligned}$$

(c) Let $Q = X_2/X_1$. For $q \geq 1$, find $P(Q \geq q)$.

10 pts

Solution. By the independence of X_1 and X_2 and the given density, the joint density is

$$f(x_1, x_2) = (3x_1^{-4})(3x_2^{-4}) = 9x_1^{-4}x_2^{-4}, \quad 1 < x_1, x_2 < \infty.$$

Hence

$$\begin{aligned} P(Q \geq q) = P(X_2 > qX_1) &= \int_{x_1=1}^\infty \int_{x_2=qx_1}^\infty 9x_1^{-4}x_2^{-4} dx_2 dx_1 = \int_{x_1=1}^\infty 9x_1^{-4} \left[\frac{x_2^{-3}}{-3}\right]_{x_2=qx_1}^\infty dx_1 \\ &= \int_{x_1=1}^\infty 3x_1^{-4}(qx_1)^{-3} dx = \frac{3}{q^3} \int_{x_1=1}^\infty x_1^{-7} dx = \boxed{\frac{1}{2q^3}} \end{aligned}$$

8. Assume the math scores on the SAT test are normally distributed with mean 500 and standard deviation 60, and the verbal scores are normally distributed with mean 450 and standard deviation 80. (In particular, under this assumption the scores take on a continuum of values and are not restricted to integer values.)

The following problems are independent of each other.

(a) Write down the density function (p.d.f.) of the math score of a randomly chosen student. (The answer should be an explicit elementary function of x, not an expression involving Φ.)

10 pts

Solution. The density of a general normal distribution with parameters μ and σ is $f(x) = (1/\sqrt{2\pi}\sigma)e^{-(x-\mu)^2/(2\sigma^2)}$. Here $\mu = 500$, $\sigma = 60$, so

$$f(x) = \boxed{\frac{1}{\sqrt{2\pi}\cdot 60}e^{-\frac{1}{2}\left(\frac{x-500}{60}\right)^2}, \quad -\infty < x < \infty.}$$

(b) Find the probability a randomly chosen student's *total* score (i.e., the sum of math and verbal scores) is between 1000 and 1100. (Assume independence of the math and verbal scores.) *Leave the answer in terms of the* Φ*-function, e.g.,* $\Phi(1100) - \Phi(1000)$.

10 pts

Solution. Let X and Y denote the math and verbal scores of the student. Then $X + Y$ is normal $N(500 + 450, 60^2 + 80^2) = N(950, 100^2)$, so

$$P(1000 < X + Y < 1100) = P\left(\frac{1000-950}{100} < \frac{X+Y-950}{100} < \frac{1100-950}{100}\right)$$
$$= \boxed{\Phi(1.5) - \Phi(0.5)}$$

Comment: For a sum of independent r.v.'s the *variances*, not the *standard deviations* add up. The standard deviation of the sum is *not* the sum of the standard deviations (i.e., not equal to $60 + 80$, or 140); to get the correct standard deviation one has to first compute the variance of the sum, $60^2 + 80^2$, then take the square root, $\sqrt{60^2 + 80^2} = 100$.

(c) Suppose two students who took both tests are chosen at random. What is the probability that the first student's math score exceeds the second student's verbal score? (Assume independence of the two scores.) *Again, leave the answer in terms of the* Φ*-function.*

10 pts

Solution. Let X and Y denote the scores of the two students. Then $X-Y$ is $N(500-450, 60^2 + 80^2) = N(50, 100^2)$, so

$$P(X > Y) = P(X - Y > 0) = P\left(\frac{X-Y-50}{100} > \frac{0-50}{100}\right) = P(Z > -1/2)$$
$$\approx 1 - \Phi(-1/2) = \boxed{\Phi(0.5)}$$

9. The following problems are independent of each other.

(a) *Using an appropriate version of normal approximation*, give an approximation for the probability of getting *between* 10 *and* 12 *heads (inclusive)* in 20 tosses with a fair coin, in terms of the Φ-function. Your answer can be left in "raw" form such as $(1/\sqrt{2\pi})\Phi((\sqrt{12} - \sqrt{10}/20)$, or $\Phi(\sqrt{12}/19.5) - \Phi(\sqrt{10}/20.5)$.

15 pts

Solution. We use the normal approximation to the binomial distribution with the 0.5 correction. The mean and standard deviation of the approximating normal distribution are $np = 20(1/2) = 10$ and $\sqrt{np(1-p)} = \sqrt{20(1/2)(1-1/2)} = \sqrt{5}$, respectively. Hence, the probability asked is

$$P(10 \le X \le 12) = \Phi\left(\frac{12.5-10}{\sqrt{5}}\right)$$
$$= \Phi\left(\frac{2.5}{\sqrt{5}}\right) - \Phi\left(\frac{-0.5}{\sqrt{5}}\right)$$
$$= \boxed{\Phi(2.5/\sqrt{5}) + \Phi(0.5/\sqrt{5}) - 1}$$

(b) The mathematics department of a large state university has funds for 50 assistantships for new graduate students. Past experience has shown that, on average, only 60 % of those applicants who are offered an assistantship accept the offer. What is the largest number of offers the Department make and still be at least 95% certain that it does not go over budget? *Note that* $\Phi(1.65) = 0.95$. *Give an approximate answer, using an appropriate approximation. The answer can be left in raw/unevaluated form (e.g., "$\lfloor \ln(\sqrt{95} - \pi) \rfloor$ offers"), but must be such that one could easily obtain a numerical answer with a simple calculator.*

15 pts

Solution. Let n be the number of offers. Using a success/failure model as in part (a) and normal approximation with parameters $\mu = np = 0.6n$ and $\sigma = \sqrt{np(1-p)} = \sqrt{0.24n}$, the probability that the Department does not go over budget becomes

$$P(\leq 50 \text{ successes in } n \text{ trials}) \approx \Phi\left(\frac{50 + 0.5 - 0.6n}{\sqrt{0.24n}}\right).$$

We set this equal to 0.95. From the normal table we get $(50.5 - 0.6n)/\sqrt{0.24n} = 1.65$, or $-0.6n - 1.65\sqrt{0.24}\sqrt{n} + 50.5 = 0$. The latter is a quadratic equation for $\sqrt{n}$ with solution

$$\sqrt{n} = \frac{1.65\sqrt{0.24} \pm \sqrt{1.65^2 0.24 + 4 \cdot 0.6 \cdot 50.5}}{-1.2}$$

Of the two roots, the one corresponding to the plus sign in $\pm$ would give a negative value for $\sqrt{n}$, so it can be eliminated. Using the other root (corresponding to the plus sign), squaring the above expression, and taking the ceiling (smallest integer greater than the given value) we get the desired answer: The number of offers the department can make and still be 95% certain that it does not go over budget is

$$n = \left\lceil \left(\frac{1}{-1.2}\left(1.65\sqrt{0.24} - \sqrt{1.65^2 0.24 + 4 \cdot 0.6 \cdot 50.5}\right)\right)^2 \right\rceil$$

(Calculating the above expression gives $n = 72$.

10. The following problems are independent of each other.

(a) Let $X_1, X_2, X_3, \ldots$ be i.i.d. random variables, with mean $\mu = \mathrm{E}(X_i)$ and variance $\sigma^2 = \mathrm{Var}(X_i)$, and let $S_n = \sum_{i=1}^n X_i$ denote the partial sums of the X_i. Given this set-up and notation, state the *Weak Law of Large Numbers* in precise mathematical form, using proper mathematical notation, and including any hypotheses/quantifiers necessary in the statement.

10 pts

Solution. The WLLN states the following: *For any $\epsilon > 0$,*

$$\lim_{n \to \infty} P\left(\left|\frac{S_n}{n} - \mu\right| \leq \epsilon\right) = 1.$$

Note: The italicized phrase is an essential part of the statement of the WLLN.

(b) If X is a random variable with mean 50 and variance 25, what can be said about the probability that X is between 40 and 60? (E.g., how large, or how small, must this probability be, given the above information?)

10 pts

Solution. [This is Example 2a in 8.2.] By Chebychev's inequality,

$$P(|X - 50| > 10) = P(|X - 50| > 5 \cdot 2) \leq \frac{1}{2^2},$$

so

$$P(40 \leq X \leq 60) = 1 - P(|X - 50| > 10) \geq 1 - \frac{1}{4} = \boxed{\frac{3}{4}}$$

11. (Extra Credit) Suppose you roll a **four-faced** fair die (i.e., a die with faces labeled $1, 2, 3, 4$, each equally likely to occur) 10 times. What is the probability that each of these four faces appears on at least one of the 10 rolls? (The result should be a simple expression, involving no more than a few terms, not a messy summation. No hints will be given on this problem.)

10 pts

Solution. Denote the outcomes of the 10 rolls as 10-tuples $(a_1, \ldots, a_{10})$ with $a_i = 1, 2, 3, 4$. Let S denote the set of all of these outcomes (i.e., tuples), and S_i denote the set of outcomes that do *not* contain the number i. The probability to compute is that of being in the complement of all of the S_i, i.e.,

$$P((S_1 \cup S_2 \cup S_3 \cup S_4)^c) = 1 - P(S_1 \cup S_2 \cup S_3 \cup S_4).$$

Now, apply inclusion/exclusion to the latter expression, to get

$$\begin{aligned} P(S) = 1 - &\sum_{1 \le i \le 4} P(S_i) \\ &+ \sum_{1 \le i < j \le 4} P(S_i S_j) \\ &- \sum_{1 \le i < j < k \le 4} P(S_i S_j S_k). \end{aligned}$$

To evaluate these terms, argue as follows:

- In the first sum, $P(S_i) = (3/4)^{10}$, and there are $\binom{4}{1}$ terms.
- In the second sum, $P(S_i S_j) = (2/4)^{10}$, and there are $\binom{4}{2}$ terms.
- In the third sum, $P(S_i S_j S_k) = (1/4)^{10}$, and there are $\binom{4}{3}$ terms.

Hence, the above becomes

$$P(S) = 1 - \binom{4}{1}(3/4)^{10} + \binom{4}{2}(2/4)^{10} - \binom{4}{3}(1/4)^{10}.$$