Midterm Review Solutions for MATH 3510

Solutions to the proof and example problems are below (in blue). In each of the example problems, the general principle is given in parentheses before the solution. If you were unable to solve a problem on the review, you might consider looking up the general principle and then trying the problem again before looking up the full solution. Note that the important thing in the proofs is the logical structure—and not the individual words used. So, don't bother trying to memorize the proofs word-for-word.

3 Basic Proofs

7. Be able to prove Rules 7.1–7.3 in Section 7.3, using only the Kolmogorov axioms.

(a) (Rule 7.1) For any finite number of mutually exclusive events $A_1, \ldots A_n$,

$$P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n).$$

Proof: We first show that $P(\emptyset) = 0$. Note that for any set A, $A \cap \emptyset = \emptyset$, so that the empty set is disjoint from any other set. In particular, the empty set is disjoint from itself, so if we let $B_i = \emptyset$ for $i = 1, 2, 3 \ldots$, then the events B_i are mutually exclusive and so by Axiom 3 $P(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i)$. Since $\bigcup_{i=1}^{\infty} \emptyset = \emptyset$, this means that $P(\emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$. That is, $P(\emptyset)$ is a number that can be added to itself infinitely many times without changing its value. The only number with this property is zero, so it must be that $P(\emptyset) = 0$. Now, given events $A_1, \ldots, A_n$ that are mutually exclusive, let $A_i = \emptyset$ for $i > n$. Note that the events $A_1, A_2, \ldots$ are mutually exclusive. So,

$$\begin{aligned} P(\bigcup_{i=1}^{n} A_i) &= P(\bigcup_{i=1}^{\infty} A_i) \\ &= \sum_{i=1}^{\infty} P(A_i) \text{ (by Axiom 3)} \\ &= \sum_{i=1}^{n} P(A_i) \text{ (since } P(A_i) = P(\emptyset) = 0 \text{ for } i > n). \end{aligned}$$

(b) (Rule 7.2) For any event A, $P(A) = 1 - P(A^C)$.
Proof: Note that $A \cup A^C = \Omega$ and $A \cap A^C = \emptyset$ (so that A and A^C are disjoint). So, by Axiom 2, $P(A \cup A^C) = P(\Omega) = 1$. Also, as A and A^C are disjoint, Rule 7.1 (proved above) tells us that $P(A \cup A^C) = P(A) + P(A^C)$. Putting this together, we have that $1 = P(A) + P(A^C)$ and subtracting $P(A^C)$ from both sides gives $P(A) = 1 - P(A^C)$.

(c) (Rule 7.3) For any events A and B, $P(A \cup B) = P(A) + P(B) - P(AB)$.
Proof: To simplify notation, let $C = AB$, $A' = A \setminus B$, and $B' = B \setminus A$. Note that A', B', and C are mutually exclusive events and that $A \cup B = A' \cup B' \cup C$, $A = A' \cup C$, and $B = B' \cup C$. As these events are mutually exclusive, Rule 7.1 gives $P(A) = P(A') + P(C)$,

$P(B) = P(B') + P(C)$, and $P(A \cup B) = P(A') + P(B') + P(C)$. Now, we compute that

$$\begin{aligned} P(A) + P(B) - P(AB) &= P(A) + P(B) - P(C) = (P(A') + P(C)) + (P(B') + P(C)) - P(C) \\ &= P(A') + P(B') + P(C) = P(A \cup B) \end{aligned}$$

as claimed.

8. Let A and B be events where the set A is a subset of the set B. Show that $P(A) \leq P(B)$.
Proof: Let $C = B \setminus A$. Since A is a subset of B, we have $B = A \cup C$. Since A and C are disjoint, we have by Rule 7.1 that $P(B) = P(A) + P(C)$. Also, by Axiom 1, $P(C) \geq 0$, so that $P(B) \geq P(A)$.

9. For any constants a and b and any random variable X, show that $\sigma^2(aX + b) = a^2\sigma^2(X)$.
Proof: We compute $\sigma^2(aX + b) = E[((aX + b) - E(aX + b))^2] = E[(aX + b - aE(X) - b)^2] = E[a^2(X - E(X))^2] = a^2E[(X - E(X))^2] = a^2\sigma^2(X)$ (where we have repeatedly used the fact that $E(aX + b) = aE(X) + b$).

10. Recall that the covariance of random variables X and Y is defined to be $\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$ and that the correlation coefficient of X and Y is defined to be $\rho(X, Y) = \frac{\text{cov}(X,Y)}{\sigma(X)\sigma(Y)}$. Suppose that $Y = aX + b$, where $a \neq 0$. Show that $\rho(X, Y)$ is either 1 or -1.
Proof: We compute

$$\begin{aligned} \text{cov}(X, Y) &= \text{cov}(X, aX + b) \\ &= E[(X - E(X))(aX + b - E(aX + b))] \\ &= E[(X - E(X))(aX + b - aE(X) - b)] \\ &= E[(X - E(X))(a(X - E(X)))] \\ &= E[a(X - E(X))^2] \\ &= aE[(X - E(X))^2] \\ &= a\sigma^2(X). \end{aligned}$$

By the previous problem, we also know that $\sigma(aX+b) = |a|\sigma(X)$. So, $\rho(X, Y) = \rho(X, aX+b) = \frac{\text{cov}(X,aX+b)}{\sigma(X)\sigma(aX+b)} = \frac{a\sigma^2(X)}{\sigma(X)(|a|\sigma(X))} = \frac{a}{|a|}$. This last quantity is equal to 1 if $a > 0$ and -1 if $a < 0$.

11. For any random variables X and Y, show that $\sigma^2(X + Y) = \sigma^2(X) + \sigma^2(Y) + 2\,\text{cov}(X, Y)$, where $\text{cov}(X, Y)$ is defined in the previous problem.
Proof: We compute

$$\begin{aligned} \sigma^2(X + Y) &= E[(X + Y - E(X + Y))^2] = E[(X + Y - E(X) - E(Y))^2] \\ &= E[((X - E(X)) + (Y - E(Y)))^2] \\ &= E[(X - E(X))^2 + 2(X - E(X))(Y - E(Y)) + (Y - E(Y))^2] \\ &= E[(X - E(X))^2] + 2E[(X - E(X))(Y - E(Y))] + E[(Y - E(Y))^2] \\ &= \sigma^2(X) + \sigma^2(Y) + 2\,\text{cov}(X, Y). \end{aligned}$$

12. Show that $\text{var}(X) = E(X^2) - E(X)^2$.
Proof: Let $\mu = E(X)$. We compute $\text{var}(X) = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] = E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2$.

13. Let X be a continuous random variable. Show that $P(a \le X \le b) = P(a < X < b)$ for any real numbers a and b with $a < b$.
Proof: Recall that X being continuous means that $P(X = c) = 0$ for all real numbers c. Then, by Rule 7.1, $P(a \le X \le b) = P(X = a) + P(a < X < b) + P(X = b) = 0 + P(a < X < b) + 0 = P(a < X < b)$.

4 Example Problems

14. The life of the battery of a certain laptop (between charges) is normally distributed with a mean of 20 hours and a standard deviation of 3 hours. What is the probability that the battery of the laptop will last for between 14 and 23 hours?
(Normal Distribution) Let X be the lifetime of the battery. Then $P(14 \le X \le 23) = P(\frac{14-20}{3} \le \frac{X-20}{3} \le \frac{23-20}{3}) = P(-2 \le Z \le 1) = \Phi(1) - \Phi(-2)$.

15. The event A has a probability of $\frac{2}{3}$ and there is a probability of $\frac{3}{4}$ that at least one of the events A and B occurs. What are the smallest and largest possible values for the probability of event B?
(Rule 7.3) We have $P(A \cup B) = P(A) + P(B) - P(AB)$, or $\frac{3}{4} = \frac{2}{3} + P(B) - P(AB)$, and $P(B) = \frac{1}{12} + P(AB)$. By Axiom 1, $P(AB) \ge 0$. By Problem #8 above, $P(AB) \le P(A) = \frac{2}{3}$. So, $P(B)$ between $\frac{1}{12}$ and $\frac{2}{3} + \frac{1}{12} = \frac{3}{4}$.

16. A military early-warning installation is constructed in a desert. The installation consists of five main detectors and a number of reserve detectors. If fewer than five detectors are working, the installation ceases to function. Every two months an inspection of the installation is mounted and at that time all detectors are replaced by new ones. There is a probability of 0.05 that any given detector will cease to function during the period between inspections. The detectors function independently of one another. How many reserve detectors are needed to ensure a probability of less than 0.1% that the system will cease to function between inspections?
(Binomial distribution) Suppose there are n detectors. Let X be the number of detectors that fail. Then we want to choose n so that $P(X \ge n-4) < 0.001$ Note that $X \sim B(n, 0.05)$, so

$$\begin{aligned} P(X \ge n-4) &= P(X = n-4) + P(X = n-3) + P(X = n-2) + P(X = n-1) + P(X = n) \\ &= \sum_{k=n-4}^{n} \binom{n}{k} (0.05)^k (0.95)^{(n-k)}. \end{aligned}$$

Explicitly computing this for small $n \ge 5$, we find that $n = 8$ is the smallest choice for which the probability is < 0.001 (in fact, for $n = 8$, the probability is 0.000371751). So, in addition to the five main detectors, they should install three reserve detectors.

17. Jean claims to have flipped a fair coin 1000 times and gotten 427 heads. Let X be the number of heads out of 1000 flips. Explain why X is approximately normal and calculate the probability of getting 427 or fewer heads out of 1000 flips. Statistically speaking, do you have reason to

doubt Jean's claim?
(Central Limit Theorem) Note that $X \sim B(1000, 0.5)$. Since $1000 \cdot 0.5 \geq 5$ and $1000 \cdot (1 - 0.5) \geq 5$, we can approximate X by a normal with mean $1000 \cdot \frac{1}{2} = 500$ and standard deviation $\frac{1}{2}\sqrt{1000} \approx 15.81$ (see p. 167). Then, $P(X \leq 427) = P(\frac{X-500}{15.81} \leq \frac{427-500}{15.81}) \approx P(Z \leq -4.62) = \Phi(-4.62)$. Even without a calculator, we know that this probability is very small, as 427 is more than four standard deviations away from the mean.

18. Three people each write down the numbers $1, \ldots, 10$ in a random order. Calculate a Poisson approximation for the probability that the three people all have at least one number in the same position.
(Poisson approximation) There are $n = 10$ trials. For a success in the ith trial, all three need to have the same number in the ith position. Whatever number the first person has in this position, there is a $\frac{1}{10}$ chance that the second person has this same number in this position and a $\frac{1}{10}$ chance that the third person does. So, there is a $\frac{1}{10^2}$ chance that both the second and the third person have this number in this position (as they write the lists of numbers independently of each other). There is a weak dependence (since a number in one position cannot appear in a different position on the same list), so we'll use a Poisson approximation X with $\lambda = np = 10 \cdot \frac{1}{100} = \frac{1}{10}$. (As $\lambda < 25$, we can't approximate this by a normal distribution.) Then, the probability that all three will have at least one number in the same position is $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-\frac{1}{10}}$.

19. Each year in Houndsville, an average of 81 letter carriers are bitten by dogs. In the past year, 117 such incidents were reported. Is this number exceptionally high?
(Poisson distribution and Central Limit Theorem) Note that we don't know n or p, but we do know $\lambda = np = 81$, so we can approximate this using a Poisson distribution X. And, since $\lambda \geq 25$, we can approximate the Poisson distribution by $N(\lambda, \lambda)$ (see p. 169–170). So, $P(X \geq 117) = P(\frac{X-81}{\sqrt{81}} \geq \frac{117-81}{\sqrt{81}}) \approx P(Z \geq 4) = 1 - \Phi(4)$. As this is four standard deviations from the mean, it is an exceptionally high number.

20. Mary and Norman have a joint checking account with \$1,000 in it. If Mary writes a check for between \$0 and \$$m$ dollars with $0 < m < 1000$ and with each amount equally likely and Norman writes a check for between \$0 and \$$n$ dollars with $0 < n < 1000$ and with each amount equally likely, what is the probability that Mary and Norman will overdraw their account?
(Probability as area) Let $\Omega = \{(x, y) \colon 0 \leq x \leq m, 0 \leq y \leq n\}$. Note that Ω is a rectangle with area mn. Let A be the event that $x + y \geq 1000$. Then A is a triangle in the upper-right hand corner of Ω (unless it happens that $m + n < 1000$, in which case A is the empty set). If $m + n < 1000$, then the probability is $P(A) = 0$. Otherwise, note that the triangle has base and height both equal to $m + n - 1000$, so that the probability is $P(A) = \frac{(m+n-1000)^2}{2mn}$.

21. The keeper of a certain king's treasure receives the task of filling each of 100 urns with 100 gold coins. While fulfilling this task, he substitutes one lead coin for one gold coin in each urn. The king suspects deceit on the part of the sentry and has two methods at his disposal of auditing the contents of the urns. The first method consists of randomly choosing one coin from each of the 100 urns. The second method consists of randomly choosing four coins from each one of 25 of the 100 urns. Which method provides the largest probability of uncovering the deceit?
(Binomial distribution). Let X_1 be the number of lead coins found using the first method and

X_2 be the number of lead coins found using the second method. Then $X_1 \sim B(100, \frac{1}{100})$ and $X_2 \sim B(25, \frac{4}{100})$. (Note that both have $np = 1$, so that they'd look the same if we did a Poisson approximation. So, this is one case where we definitely don't want a Poisson approximation.) Then

$$P(X_i \geq 1) = 1 - P(X_i = 0) = 1 - \binom{n}{0} p^0 (1-p)^n \text{ for } i = 1, 2.$$

This gives $P(X_1 \geq 1) \approx 0.634$ and $P(X_2 \geq 1) \approx 0.640$, so that the second method is slightly more likely to uncover the deceit. (Note that these percentages are quite close, which makes sense as they can both by approximated by the same Poisson distribution. The Poisson approximation of this probability is 0.632.)

22. A dart is thrown at random on a rectangular board. The board measures 20 cm by 50 cm. A hit occurs if the dart lands within 5 cm of any of the four corner points of the board. What is the probability of a hit?
(Probability as area) Let $\Omega = \{(x, y), 0 \leq x \leq 20, 0 \leq y \leq 50\}$. Note that Ω is a rectangle with area 1000. The four corners are each squares with area 25, so the probability of hitting one of the four corner regions is $\frac{4 \cdot 25}{1000} = \frac{1}{10}$.

23. Calculate a Poisson approximation for the probability that two consecutive numbers will appear in a lotto drawing of six numbers from the numbers $1, \ldots, 45$.
(Poisson approximation) Consider 44 Bernoulli trials for the pairs of consecutive numbers $(1, 2), \ldots, (44, 45)$. In each trial, the probability that both of these numbers will appear in the lotto drawing is $\frac{\binom{43}{4}}{\binom{45}{6}}$. Note that there is a weak dependence between the trials (as a failure in the $(1, 2)$ trial, for example, means that the numbers 1 and 2 also can't both appear in any of the following trials), so we want to use a Poisson approximation. Let $\lambda = 44 \cdot \frac{\binom{43}{4}}{\binom{45}{6}}$ and let $X \sim \text{Pois}(\lambda)$. Then $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-\lambda} \approx 0.487$.

24. An integer is chosen at random from the numbers $1, \ldots, 1000$. What is the probability that the chosen integer is divisible by at least one of 2, 3, and 5. What is the probability that the chosen integer is divisible by none of 2, 3, and 5?
(Inclusion-Exclusion) Let A be the event that the number is divisible by 2, B be the event that the integer is divisible by 3, and C be the event that the integer is divisible by 5. Then $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC) = \frac{500}{1000} + \frac{333}{1000} + \frac{200}{1000} - \frac{166}{1000} - \frac{100}{1000} - \frac{66}{1000} + \frac{33}{1000} = \frac{734}{1000}$. The probability that the number is divisible by none of 2, 3, and 5 is $P((A \cup B \cup C)^C) = 1 - \frac{734}{1000} = \frac{266}{1000}$.