You are not allowed to use books or notes. Calculators approved for the actuarial exams are permitted. The last page contains a table for the cumulative distribution function of a standard normal variable. Please read the directions carefully. For the free response problems, show your work. The maximum obtainable score is 400. The exam will be graded out of 350 points, so you may earn extra credit. You have 2 hours to complete it. Do not detach any pages except for the table. Scratch paper will be provided on request.
1. Please clearly mark only one answer for each of the multiple choice questions. If the exact answer is not available, choose the closest numerical answer.

(a) (15 points) Suppose $A, B,$ and $C$ are events such that $P(A \cup B) = 0.55, P(A \cup C) = 0.6, P(A \cup (B \cap C)) = 0.4,$ and $P(A) = 0.2$. What is $P(A \cup B \cup C)$?

<table>
<thead>
<tr>
<th></th>
<th>A 0.65</th>
<th>B 0.7</th>
<th>C 0.75</th>
<th>D 0.8</th>
<th>E 0.85</th>
</tr>
</thead>
</table>

**Solution:** Using inclusion exclusion principle,

$$P(A \cup B \cup C) = P((A \cup B) \cup (A \cup C)) = P(A \cup B) + P(A \cup C) - P((A \cup B) \cap (A \cup C)) = P(A \cup B) + P(A \cup C) - P(A \cup (B \cap C)) = 0.55 + 0.6 - 0.4 = 0.75.$$ 

The answer is **C**.

(b) (15 points) A coin lands heads with probability 0.4. A coin is flipped repeatedly. Let $p_1$ denote the probability of showing exactly 3 heads in the first 15 flips, $p_2$ the probability of showing the 1-st heads on the 5-th roll, and $p_3$ of showing the 2-nd heads on the 8-th roll, respectively. What is the correct ordering of $p_1, p_2,$ and $p_3$?

<table>
<thead>
<tr>
<th></th>
<th>A $p_1 &gt; p_2 &gt; p_3$</th>
<th>B $p_1 &gt; p_3 &gt; p_2$</th>
<th>C $p_2 &gt; p_1 &gt; p_3$</th>
<th>D $p_2 &gt; p_3 &gt; p_1$</th>
<th>E $p_3 &gt; p_1 &gt; p_2$</th>
</tr>
</thead>
</table>

**Solution:** The number of heads in 15 flips is a binomial random variable with parameters $n = 15, p = 0.4$, so $p_1 = \binom{15}{3}0.4^3 \times 0.6^{12} \approx 0.0634$. The index of the first flip resulting in heads is a geometric random variable (with the parameter $p = 0.4$), so $p_2 = 0.6^{5-1} \times 0.4 \approx 0.0518$. The index of the second heads flip is a negative binomial random variable (with $p = 0.4$ and $r = 2$). Alternatively, to get the second heads on flip $k$ implies that the first $k - 1$ flips have exactly 1 heads (probability $(k^{-1})0.4 \times 0.6^{k-2}$), and the $k$-th flip is heads (probability 0.4), so $p_3 = \binom{8}{1}0.4 \times 0.6^{8-2} \times 0.4 \approx 0.0523$. Thus $p_1 > p_3 > p_2$, and the answer is **B**.

(c) (15 points) A fair coin is flipped repeatedly until it shows 100 heads. What is the probability that the number of needed flips is no more than 200? (Hint: What is the random variable in this problem? Use continuity correction.)

<table>
<thead>
<tr>
<th></th>
<th>A 0.50</th>
<th>B 0.52</th>
<th>C 0.54</th>
<th>D 0.56</th>
<th>E 0.58</th>
</tr>
</thead>
</table>

Solution: Let \( X \) denote the number of flips needed to get 100 heads. Let \( X_i \) denote the number of flips need to get \( i \)-th head after \( i-1 \)-th head appeared. Then \( X = \sum_{i=1}^{100} X_i \). \( X_i \)'s are independently and identically distributed geometric random variables with mean \( E[X_i] = 2 \) and variance \( \text{Var}(X_i) = 2 \). Applying the Central Limit Theorem with continuity correction,

\[
P(X \leq 200) = P(X \leq 200.5) = P\left( \sum_{i=1}^{100} X_i \leq 200.5 \right) \\
= P\left( \frac{\sum_{i=1}^{100} -100 \times 2}{\sqrt{100 \times 2}} \leq \frac{200.5 - 100 \times 2}{\sqrt{100 \times 2}} \right) \approx \Phi \left( \frac{200.5 - 100 \times 2}{\sqrt{100 \times 2}} \right) \\
\approx \Phi (0.04) = 0.5160.
\]

The answer is B.

(d) (15 points) A five card poker hand is dealt. Two of the five cards are uncovered; they are an ace and a king. What is the probability that this hand is a three of a kind? A three of a kind is a five-card hand that has three cards of the same denomination and two cards each of a denomination different from the rest of the cards in the hand. (Hint: consider both the case when the three are either aces or kings and the case when the three are in another denomination.)

\[
A \ 0.010 \quad B \ 0.013 \quad C \ 0.016 \quad D \ 0.020 \quad E \ 0.025
\]

Solution: There are 50 unopen cards, and there are three cards to open, so discounting the order of the cards, there are \( \binom{50}{3} \) possible card combinations for the unopened cards, all equally likely. There are three possibilities to complete a three of a kind from a king and an ace: have two more aces and another card which is not a king or an ace, have two more kings and another card which is not a king or an ace, or have three cards in a denomination other than a king or an ace. There are \( \binom{3}{2} \) ways to choose two aces from three remaining aces, and \( \binom{52-8}{1} \) ways to choose one card other than a king or an ace, so the number of combinations of two aces and a non-ace, non-king card is \( \binom{3}{2} \times \binom{44}{1} = 132 \). The number of completions to get three of a kind with kings is the same, 132. To get a three of a kind with another denomination, we need to pick one of other 11 denominations, and then to pick three cards (out of four) in that denomination, a total \( \binom{11}{1} \times \binom{4}{3} = 44 \) ways. Thus the probability of completing a three of a kind is \( (2 \times 132 + 44) / \binom{50}{3} \approx 0.0157 \). The answer is C.

(e) (15 points) A fair coin is flipped 7200 times, and a fair die is rolled 1000 times. What is the probability that the number of heads is larger than the sum of all die roll values?

\[
A \ 0.81 \quad B \ 0.84 \quad C \ 0.88 \quad D \ 0.93 \quad E \ 0.95
\]
Solution: Let $X$ be the number of heads, and let $Y$ be the sum of all die rolls. $X$ is a binomial random variable, and its distribution can be approximated with a normal distribution with mean $7200 \times \frac{1}{2} = 3600$ and variance $7200 \times \frac{1}{2} \times \left(1 - \frac{1}{2} \right) = 1800$. $Y$ is a sum of values of independent rolls (mean $3.5 \times 1000 = 3500$ and variance $1000 \times \frac{35}{12}$), and its distribution can be approximated (using the Central Limit Theorem) with a normal with mean $3.5 \times 1000 = 3500$ and variance $1000 \times \frac{35}{12}$. The sum of two normal random variables is a normal random variable with mean equal to the sum of the means, and variance equal to the sum of the variances, so $X - Y$ is approximately normal with mean $3600 - 3500 = 100$ and variance $1800 + 1000 \times \frac{35}{12} \approx 4716.67$. Using the continuity correction,

$$P(X > Y) = P(X - Y > 0) = P(X - Y \geq 0.5) = P\left(\frac{X - Y - 100}{\sqrt{4716.67}} \geq \frac{0.5 - 100}{\sqrt{4716.67}}\right) \approx 1 - \Phi\left(\frac{0.5 - 100}{\sqrt{4716.67}}\right) = \Phi\left(\frac{100 - 0.5}{\sqrt{4716.67}}\right) = \Phi(1.45) = 0.9265.$$ 

the answer is D.

(f) (15 points) A coin is flipped repeatedly until the first heads appears. What is the probability of getting heads on any given flip if the probability that the coin is going to be flipped more than twice is 0.64?

Let $p$ be the probability of getting heads on any flip. The index of the first heads flip, $X$, is a geometric random variable with the probability mass function $p(k) = P(X = k) = (1 - p)^{k-1} p$, $k = 1, 2, \ldots$.

$$0.64 = P(X > 2) = 1 - P(X = 0) - P(X = 1) = 1 - p - (1 - p)p = p^2 - 2p + 1.$$ 

Solving the corresponding quadratic equation $p^2 - 2p + 0.36$, we find roots $p = 0.2$ and $p = 1.8$. Since the probability of a coin landing heads cannot be larger than 1, $p = 0.2$. The answer is D.

(g) (15 points) A random variable $X$ has a probability density function

$$f(x) = \begin{cases} ae^{-2x} + bx^2e^{-x} & x \geq 0, \\ 0 & \text{otherwise}. \end{cases}$$ 

What is the value of $a$ if $E[X] = \frac{1}{2}$?

$$A \frac{1}{3} \quad B \ 0.75 \quad C \ 1 \quad D \ 1.5 \quad E \ 2$$
The total density has to integrate to 1, so
\[ 1 = \int_{0}^{\infty} ae^{-2x} + bx^2 e^{-x} \, dx = a \int_{0}^{\infty} e^{-2x} \, dx + b \int_{0}^{\infty} x^2 e^{-x} \, dx = \frac{a}{2} + \Gamma(3) b; \]
\[ E[X] = \int_{0}^{\infty} x (ae^{-2x} + bx^2 e^{-x}) \, dx = a \int_{0}^{\infty} xe^{-2x} \, dx + b \int_{0}^{\infty} x^3 e^{-x} \, dx = \frac{a}{4} + \Gamma(4) b. \]

From the first equation, \( b = \frac{1}{2} - \frac{a}{4}. \) Plugging it into the second equation,
\[ \frac{1}{2} = a \left( \frac{1}{2} - \frac{a}{4} \right) = 3 - \frac{5a}{4} \implies a = 2. \] (Also \( b = 0. \))

(h) Batteries produced by a manufacturer A last on average 200 hours while batteries made by B last on average 180 hours. A flashlight has had a new battery put in, with probability 0.6 that it is made by A, and probability 0.4 that it is made by B. Assuming that the battery life for each manufacturer is exponentially distributed, what is the standard deviation for the life of the battery in the flashlight?

\[ A \ 179.50 \quad B \ 185.13 \quad C \ 192 \quad D \ 192.25 \quad E \ 192.50 \]

Solution: Let \( Y \) be a random variable indicating whether the battery manufacturer is A, \( P(Y = 1) = 0.6, \) \( P(Y = 0) = 0.4. \) Let \( X \) denote the life of a battery, so both \( X|Y = 1 \) and \( X|Y = 0 \) are exponentially distributed, with \( E[X|Y = 1] = 200 \) (and therefore \( \text{Var}(X|Y = 1) = 200^2 = 40000 \)) and \( E[X|Y = 0] = 180 \) \( \text{Var}(X) = 180^2 = 32400 \). Then
\[ \text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E[X|Y]); \]
\[ E[\text{Var}(X|Y)] = 0.6 \times 40000 + 0.4 \times 32400 = 36960; \]
\[ E[E[X|Y]] = 0.6 \times 200 + 0.4 \times 180 = 192; \]
\[ \text{Var}(E[X|Y]) = E[(E[X|Y] - E[E[X|Y]])^2] \]
\[ = 0.6 \times (200 - 192)^2 + 0.4 \times (180 - 192)^2 = 96; \]
\[ \text{Var}(X) = 36960 + 96 = 37056; \]
\[ \text{SD}(X) = \sqrt{\text{Var}(X)} \approx 192.50. \]

The answer is E.

(i) Two sprinters run 100m race. The first sprinter finish time is uniformly distributed between 9.7 and 9.95 seconds, while the other sprinters finish time is uniformly distributed between 9.8 and 9.9. Assuming that that the sprinters run independently of each other, what is the probability that they will run the race within 0.01 seconds (photo-finish)?

\[ A \ 0.01 \quad B \ 0.02 \quad C \ 0.04 \quad D \ 0.08 \quad E \ 0.16 \]
Solution: Let \( X \) and \( Y \) denote the finish times for the first and the second sprinter, respectively. We are asked to find \( P( |X-Y| < 0.01) \). The joint probability density function for \( X \) and \( Y \) is

\[
f_{XY}(x,y) = f_X(x)f_Y(y) = \begin{cases} 
\frac{1}{9.95-9.7} \times \frac{1}{9.9-9.8} = 40 & x \in (9.7, 9.95), y \in (9.8, 9.9), \\
0 & \text{otherwise.}
\end{cases}
\]

Then

\[
P( |X-Y| < 0.01) = \int_{|x-y|<0.01} \int_{9.8}^{9.9} f_{XY}(x,y) \, dy \, dx = \int_{9.8}^{9.9} 40 \, dy \times 0.02 = 40 \times 0.02 \times 0.1 = 0.08.
\]

The answer is D.

(j) (15 points) The length of a person’s step is uniformly distributed between 50 and 100 cm. What is the probability that the person would walk less than 740m in 1000 steps? (Assuming that the lengths of the steps are independent of one another.)

\[
A \ 0.14 \quad B \ 0.17 \quad C \ 0.20 \quad D \ 0.23 \quad E \ 0.26
\]

Solution: Denote by \( X \) the total distance in 1000 steps, and let \( X_i \) denote the length of the \( i \)-th step (in meters), so \( X = \sum_{i=1}^{1000} X_i \). Let \( \mu = E[X_i] = \frac{0.5+0.1}{2} = 0.75 \), and \( \sigma^2 = \text{Var}(X_i) = \frac{(1-0.5)^2}{12} = \frac{1}{48} \). Then using the Central Limit Theorem,

\[
P(X \leq 740) = P \left( \frac{X - 740 \times 0.75}{\sqrt{1000 \times \frac{1}{48}}} < \frac{740 - 1000 \times 0.75}{\sqrt{1000 \times \frac{1}{48}}} \right) \\
\approx \Phi \left( \frac{740 - 1000 \times 0.75}{\sqrt{1000 \times \frac{1}{48}}} \right) \approx \Phi (-1.10) \\
= 1 - \Phi (1.10) = 1 - 0.8643 = 0.1357.
\]

The answer is A.
2. (20 points) Two urns contain identical red and blue balls. The first urn has 5 of each while the second contains 15 red balls and an unknown number of blue balls. The first urn is chosen with probability 0.4, the second with probability 0.6, and one ball is drawn from the chosen urn. If the probability of drawing a blue ball is 0.35, how many blue balls are in the second urn?

Solution: Let $U_1$ denote the event that the first urn is chosen, and let $U_2$ denote the event that the second urn is chosen. ($U_2 = U_1^c$.) Let $B$ denote the event that a blue ball is drawn. Let $n$ denote the number of blue balls in the second urn. Then

$$0.35 = P(B) = P(U_1)P(B|U_1) + P(U_2)P(B|U_2) = 0.4 \times 0.5 + 0.6 \times \frac{n}{15 + n};$$

$$0.15 = 0.6 \times \frac{n}{15 + n} \Rightarrow n = 5.$$

3. (20 points) Three different bacteria types, A, B, and C, grow one type per plate. The number of colonies for a plate of type A is a Poisson random variable with mean 3; for type B, it is a binomial random variable with $n = 10$, $p = 0.3$; for type C, it is a geometric random variable with mean 3. A plate is picked at random (with equal probability for all bacteria types). If there are three colonies on a plate, find the probabilities that the chosen plate is of type A, type B, and type C.

Solution: Let $A$, $B$, and $C$ denote the events of picking a plate with bacteria of type A, B, and C, respectively. Then $P(A) = P(B) = P(C) = \frac{1}{3}$. Let $X$ denote the number of bacterium colonies on the chosen plate. $X|A$ is a Poisson random variable with $\lambda = 3$, $X|B$ is a binomial random variable with $n = 10$, $p = 0.3$, and $X|C$ is a geometric random variable with $p = \frac{1}{3}$. Using Bayes’ Rule,

$$P(A|X = 3) = \frac{P(A)P(X = 3|A)}{P(X = 3)} = \frac{\frac{1}{3}e^{-3}3^3}{\frac{1}{3}e^{-3}3^3 + \frac{1}{3}(\frac{10}{3})0.3^3 \times 0.7^7 + \frac{1}{3} \times \left(\frac{2}{3}\right)^2 \times \frac{1}{3}} \approx 0.3506;$$

$$P(B|X = 3) = \frac{P(B)P(X = 3|B)}{P(A)P(X = 3|A) + P(B)P(X = 3|B) + P(C)P(X = 3|C)} \approx 0.4176;$$

$$P(C|X = 3) = \frac{P(C)P(X = 3|C)}{P(A)P(X = 3|A) + P(B)P(X = 3|B) + P(C)P(X = 3|C)} \approx 0.2318.$$

4. 

$$f_{XY}(x, y) = \begin{cases} 24xy & x \geq 0, \ y \geq 0, \ x + y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$
(a) (10 points) Find \( f_X(x) \) and \( f_Y(y) \)

**Solution:**

\[
\begin{align*}
f_X(x) & = \int_0^{1-x} 24xy dy = 12x y^2 \bigg|_0^{1-x} = 12x (1-x)^2, \quad x \in [0, 1]. \\
\end{align*}
\]

By symmetry, \( f_Y(y) = 12y (1-y)^2, \quad y \in [0, 1]. \)

(b) (15 points) Find \( E[X], E[Y], \) Var \((X)\), and Var \((Y)\).

**Solution:**

\[
\begin{align*}
E[X] & = \int_0^1 x \times 12x (1-x)^2 \, dx = \int_0^1 12 \left( x^2 - 2x^3 + x^4 \right) \, dx \\
& = \left( 4x^3 - 6x^4 + \frac{12}{5} x^5 \right) \bigg|_0^1 = 0.4; \\
\text{Var}(X) & = E[X^2] - (E[X])^2 = \int_0^1 12 \left( x^3 - 2x^4 + x^5 \right) \, dx - 0.16 \\
& = \left( 3x^4 - \frac{24}{5} x^5 + 2x^6 \right) \bigg|_0^1 = 0.2 - 0.16 = 0.04.
\end{align*}
\]

By symmetry, \( E[Y] = 0.4, \) and \( \text{Var}(Y) = 0.04. \)

(c) (15 points) Find \( \text{Cov}(X,Y) \) and \( \rho(X,Y) \).

**Solution:**

\[
\begin{align*}
\text{Cov}(X,Y) & = E[XY] - E[X] E[Y] = \int_0^1 \int_0^{1-x} 24xy^2 \, dy \, dx - 0.16 \\
& = \int_0^1 8x^2 y^3 \bigg|_0^{1-x} \, dx - 0.16 = \int_0^1 8x^2 (1-x)^3 \, dx - \frac{4}{25} = \frac{2}{15} - \frac{4}{25} \\
& = -\frac{2}{75}; \\
\rho(X,Y) & = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} = \frac{-\frac{2}{75}}{\frac{1}{25}} = -\frac{2}{3}.
\end{align*}
\]

5. Let \( X \) be a uniform random variable on \((0,3)\), and let \( Y \) be an exponential random variable with mean 2. Assume that \( X \) and \( Y \) are independent.

(a) (15 points) Find the mean and the variance of \( X + Y \).

**Solution:** \( E[X] = \frac{1}{2} (0 + 3) = 1.5, \) \( E[Y] = 2. \) \( E[X+Y] = E[X] + E[Y] = 1.5 + 2 = 3.5. \) \( \text{Var}(X) = \frac{3^2}{12} = \frac{3}{4}, \) \( \text{Var}(Y) = 4, \) \( \text{Var}(X+Y) = \text{Var}(X)+\text{Var}(Y) = \frac{19}{4}. \)

(b) (15 points) Find the moment generating function for \( X + Y \).
Solution: For independent random variables \( X \) and \( Y \), \( M_{X+Y} (t) = M_X (t) M_Y (t) \).

\[
M_X (t) = E [e^{tX}] = \int_0^\infty 3 e^{tx} \frac{1}{3} e^{tx} = \frac{1}{3t} (e^{3t} - 1);
\]

\[
M_Y (t) = E [e^{tY}] = \int_0^\infty 2 e^{-\frac{1}{2}y} e^{ty} = \int_0^\infty 2 e^{(t-\frac{1}{2})y} dy = \frac{1}{2t - 1};
\]

\[
M_{X+Y} (t) = M_X (t) M_Y (t) = \frac{e^{3t} - 1}{3t (2t - 1)}.
\]

6. (20 points) 50 distinct numbers are chosen at random from \( \{1, \ldots, 100\} \). What is the expected number of pairs of consecutive numbers among these 50 numbers? (For example, \( \{1, 2, 3\} \) contains two pairs of consecutive numbers, \{1, 2\}, and \{2, 3\}).

Solution: There are 99 possible pairs, with the smaller numbers in the pair \( k = 1, \ldots, 99 \). Let \( X_k \) be an indicator random variable with \( X_k = 1 \) if both \( k \) and \( k + 1 \) were chosen, and 0 otherwise. Then the number of chosen pairs \( X = \sum_{k=1}^{99} X_k \).

\[
E [X] = \sum_{k=1}^{99} E [X_k] = \sum_{k=1}^{99} P (\text{both } k \text{ and } k + 1 \text{ were chosen}) = \sum_{k=1}^{99} \frac{\binom{98}{48}}{\binom{100}{50}}
\]

\[
= \frac{99 \times 98! \times (50!)^2}{48!50!100!} = \frac{50 \times 49}{100} = 24.5.
\]

7. Let \( X \) be a uniform random variable over \( (\frac{1}{2}, 2) \), and \( Y \) be an exponential random variable with mean \( \frac{1}{2} \). Assume \( X \) and \( Y \) are independent.

(a) (10 points) What is the joint probability density function for \( X \) and \( Y \)?

Solution:

\[
f_X (x) = \begin{cases} \frac{1}{2} & x \in (\frac{1}{2}, 2), \\ 0 & \text{otherwise}, \end{cases}
\]

\[
f_Y (y) = \begin{cases} 2e^{-2y} & y \geq 0, \\ 0 & \text{otherwise}; \end{cases}
\]

\[
f_{X,Y} (x, y) = f_X (x) f_Y (y) = \begin{cases} \frac{4}{3} e^{-2y} & x \in (0.5, 2), y \geq 0, \\ 0 & \text{otherwise}. \end{cases}
\]

(b) (15 points) Find \( E [X^2 Y^2] \).

Solution: Since \( X \) and \( Y \) are independent, \( E [g (X) h (Y)] = E [g (X)] E [h (Y)] \), so

\[
E [X^2 Y^2] = E [X^2] E [Y^2] = \int_{0.5}^{2} \frac{2x^2}{3} dx \int_0^\infty 2y^2 e^{-2y} dy
\]

\[
= \left( \frac{2x^3}{9} \right)_{0.5} \left( \frac{1}{4} + \frac{1}{4} \right) \approx 0.8873.
\]

(c) (15 points) Let \( U = XY \), and let \( V = \frac{Y}{X} \). What is the joint probability density function for \( u \) and \( v \)?
Solution: First, we express, $X$ and $Y$ as functions of $U$ and $V$:

$$UV = Y^2, \quad \frac{U}{V} = X^2,$$

so $X = \sqrt{\frac{U}{V}}$ and $Y = \sqrt{UV}$;

$$Y = \sqrt{UV} \geq 0, \quad 0.5 < X = \frac{U}{V} < 2, \text{ so } V \geq 0 \text{ and } 0.5V < U < 2V.$$  

Next, we find the Jacobian and its determinant

$$J = \left( \begin{array}{cc} \frac{\partial x(u,v)}{\partial u} & \frac{\partial x(u,v)}{\partial v} \\ \frac{\partial y(u,v)}{\partial u} & \frac{\partial y(u,v)}{\partial v} \end{array} \right) = \left( \begin{array}{cc} \frac{1}{2} \sqrt{\frac{1}{uv}} & -\frac{1}{2} \sqrt{\frac{u}{v^3}} \\ \frac{1}{2} \sqrt{\frac{u}{v^3}} & \frac{1}{2} \sqrt{\frac{v}{u}} \end{array} \right), \text{ so } |\det J| = \frac{1}{2v}.$$  

Therefore,

$$f_{UV}(u,v) = f_{XY}(x,y) |\det J| = \frac{2}{3v}e^{-\sqrt{uv}}, \quad v > 0, \quad 0.5v < u < 2v, \text{ and } 0 \text{ otherwise.}$$  

8. The mean daily temperature high in May in West Lafayette is 70 degrees Fahrenheit. The probability that a daily high is either at least 78 or at most 62 is 0.36.

(a) (15 points) What is the smallest value of the standard deviation of the daily high temperature? (Hint: use Chebyshev’s inequality.)

Solution: Let $X$ denote the daily high in May, and $\mu = E[X] = 70$. Let $\sigma = \sqrt{\text{Var}(X)}$. By Chebyshev’s inequality,

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2};$$

so

$$P(|X - \mu| \geq 8) = 0.36 \leq \frac{\sigma^2}{64} \Rightarrow \sigma^2 \geq 64 \times 0.36 \Rightarrow \sigma \geq 8 \times 0.6 = 4.8.$$  

(b) (15 points) Now assume that the daily high is normally distributed. Find the standard deviation.

Solution:

$$P(|X - \mu| \geq 8) = 1 - P\left(\frac{-8}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{8}{\sigma}\right) = 1 - \left(\Phi\left(\frac{8}{\sigma}\right) - \Phi\left(-\frac{8}{\sigma}\right)\right)$$

$$= 2 - 2\Phi\left(\frac{8}{\sigma}\right) = 0.36 \Rightarrow \Phi\left(\frac{8}{\sigma}\right) = 0.82.$$  

Looking up in the normal table, the corresponding value is between 0.91 and 0.92. (It is closer to 0.92.) Then $\frac{8}{\sigma} = 0.92$; therefore, $\sigma = \frac{8}{0.92} \approx 8.70.$
9. Extra credit problems, attempt once the rest of the problems are completed. No partial credit will be given; points are earned only for correctly solved problems or subproblems.

(a) (15 points) Let $X$ be a discrete random variable with the probability mass function

\[
p(k) = \begin{cases} 
\frac{1}{k(k+1)} & k = 1, 2, \ldots, \\
0 & \text{otherwise}.
\end{cases}
\]

What is $E[X]$?

**Solution:**

\[
E[X] = \sum_{k=1}^{\infty} p(k) k = \sum_{k=1}^{\infty} \frac{k}{k(k+1)} = \sum_{k=1}^{\infty} \frac{1}{k+1};
\]

\[
\lim_{n \to \infty} \left( \sum_{k=1}^{n} \frac{1}{k} - \log n \right) = \gamma \text{ where } \gamma \text{ is the Euler constant;}
\]

\[
E[X] = \infty.
\]

In other words, the expected value for $X$ is not defined.

(b) (15 points) 50 distinct numbers are chosen at random from $\{1, \ldots, 100\}$. What is the probability mass function for the smallest number among the chosen?

**Solution:** If $k$ is the minimum among chosen numbers, it means that none of the $1, \ldots, k-1$ were chosen, and $k$ was. Let $X$ denote the minimum of the chosen numbers,

\[
P(X = k) = P(1, \ldots, k-1 \text{ not chosen}, k \text{ is chosen})
\]

\[
= P(1, \ldots, k-1 \text{ not chosen}) P(k \text{ is chosen} | 1, \ldots, k-1 \text{ not chosen})
\]

\[
= \frac{100-(k-1)}{50} \times \frac{100-k}{49} = \frac{100-k}{49}.
\]

(c) (20 points) Let $X$ and $Y$ be exponential random variables with parameters $\lambda_x$ and $\lambda_y$, respectively, taking values only on positive real numbers. Find the distribution function for $X/(X+Y)$.

**Solution:** Let $Z = \frac{X}{X+Y}$. Note that $\frac{X}{X+Y} = \frac{1}{1+\frac{Y}{X}}$, and since $\frac{Y}{X} \in (0, \infty)$, $\frac{1}{1+\frac{Y}{X}} \in (0, 1)$, so $Z \in (0, 1)$. In other words, $F_Z(z) = 0$ if $z \leq 0$, and $F_Z(z) = 1$ if $Z \geq 1$. 
Assume \( z \in (0, 1) \).

\[
P(Z \leq z) = P\left( \frac{X}{X+Y} \leq z \right) = P(X \leq z(X+Y)) = P\left( Y \geq \frac{1-z}{z} X \right)
\]

\[
= \int\int_{y \geq \frac{1-z}{z} x} f_{XY}(x, y) \, dy \, dx = \int_0^\infty \int_{x(1-z)\over z}^\infty \lambda_x e^{-\lambda_x x} \lambda_y e^{-\lambda_y y} \, dy \, dx
\]

\[
= \int_0^\infty \lambda_x e^{-\lambda_x x} \left( -e^{-\lambda_y y} \right)^{\infty \over x(1-z)\over z} \, dx = \int_0^\infty \lambda_x e^{-\lambda_x x} e^{-\lambda_y y \left(1 \over x \right) \over z} \, dx
\]

\[
= \int_0^\infty \lambda_x e^{-x(\lambda_x + \lambda_y {1-z \over z})} \, dx = \frac{\lambda_x}{\lambda_x + \lambda_y {1-z \over z}} = \frac{\lambda_x z}{(\lambda_x - \lambda_y) z + \lambda_y};
\]

\[
F_Z(z) = \begin{cases} 
0 & z \leq 0, \\
{\lambda_x z \over (\lambda_x - \lambda_y) z + \lambda_y} & 0 < z < 1, \\
1 & z \geq 1.
\end{cases}
\]