No books, no notes, only SOA-approved calculators. You must show work, unless the question is a true/false, multiple-choice, or fill-in-the-blank question. Numerical answers should be given to 3 places, e.g. 97.9%.

1. (6 points) Suppose we have three boxes:

   - a box containing two gold coins,
   - a box containing two silver coins,
   - a box containing one gold coin and a silver coin.

   We choose a box at random and withdraw one coin at random. It happens to be a gold coin. What is the probability that the other coin in the same box is also a gold coin? Select the option below closest to the correct answer.

   1. ____________

   A. 0.63
   B. 0.56
   C. 0.46
   D. 0.31

   Solution: This is the Bertrand Box Puzzle (1889), an early precursor of the Monty Hall Problem. The probability that we have selected a box with two coins of the same type is 2/3, so the answer is 2/3.

2. (6 points) A student answers a multiple-choice examination question that offers four possible answers. Suppose the probability that the student knows the answer to the question is 0.8 and the probability that the student will guess is 0.2. Assume that if the student guesses, the probability of selecting the correct answer is 0.25. If the student correctly answers a question, what is the probability that the student really knew the correct answer?

   2. ____________

   Solution: This is problem 2.133. Define the events: G: student guesses, C: student is correct. Apply Bayes' Formula:

$$P(G \mid C) = \frac{P(C \mid G)P(G)}{P(C \mid G)P(G) + P(C \mid G)P(G)} = \frac{1(.8)}{1(.8) + .25(.2)} = 0.9412.$$

3. (6 points) As items come to the end of a production line, an inspector chooses which items are to go through a complete inspection. Ten percent of all items produced are defective. Sixty percent of all defective items go through a complete inspection, and 20% of all good items go through a complete inspection. Given that an item is completely inspected, what is the probability it is defective?

3. ____________

Solution: This is problem 2.173. Let D be the event that an item is defective. Let C be the event that an item goes through a complete inspection. Thus $P(D) = 0.1$, $P(C \mid D) = 0.6$, and $P(C \mid \bar{D}) = 0.2$. Thus,

$$P(D \mid C) = \frac{P(C \mid D)P(D)}{P(C \mid D)P(D) + P(C \mid \bar{D})P(\bar{D})} = .25.$$

4. (6 points) Recall the original Monty Hall Problem: the game show has three curtains, one of which conceals a car, and the other two of which conceal goats. The contestant chooses a curtain at random, and the host (who knows where the car is) opens another curtain to display a goat. The contestant is offered the opportunity to switch from his original choice to the remaining unopened curtain.

   Suppose now the host does not always open another curtain and offer the option to switch, but chooses each time whether to do so after seeing the contestant's choice of curtain. Assume the host wishes to minimize the contestant's chances of winning the car. Which of the following statements is most accurate?

4. ____________

   A. The contestant should switch curtains whenever offered a choice and his chances of winning are unchanged.

   B. With the correct strategy for offering choices, the host can reduce the contestant's chances of winning the car to less than 1/3.

   C. The contestant should switch curtains whenever offered a choice, but his chances of winning are not the same as in the original problem.

   D. None of the above statements is accurate.

Solution: Consider the following strategy on the part of the host: show a goat and offer a choice if, and only if, the contestant has selected the curtain hiding the car. Then the contestant will lose every time he elects to switch. The contestant can keep his chances at 1/3 by refusing to ever switch curtains.

5. Suppose that Y is normally distributed with mean μ and variance σ^2 and that $U = e^Y$.

(a) (6 points) Fill in the blanks to complete the statements about the relationship between the CDFs F_Y and F_U.

$$
\begin{aligned}
F_U(u) &= P(\underline{U \le u}) && \text{by definition of CDF} \\
&= P(\underline{e^Y} \le u) && \text{by definition of } U \\
&= P(Y \le \underline{\log u}) && \\
&= \underline{F_Y(\log u)} && \text{by definition of } F_Y
\end{aligned}
$$

(b) (6 points) Find the density function $f_U(u)$. *Hint:* Your answer should depend on μ and σ.

Solution: Differentiate the last line above to get

$$
f_U(u) = \begin{cases} \dfrac{1}{u\sqrt{2\pi}} e^{\frac{(\log u - \mu)^2}{2\sigma^2}} & \text{if } u \ge 0 \\ 0 & \text{if } u < 0 \end{cases}
$$

(c) (3 points) Fill in the blank: the definition of the MGF $m_Y(t)$ is $E[\underline{e^{tY}}]$.

(d) (3 points) Use the moment-generating function of Y to find $E[U]$. *Hint:* Your answer should depend on μ and σ.

(d) ______________

Solution:

$$E[U] = E[e^Y] = E[e^{1Y}] = m_Y(1) = e^{\mu t + \frac{\sigma^2 t^2}{2}}|_{t=1} = e^{\mu + \frac{\sigma^2}{2}}$$

(e) (3 points) Use the moment-generating function of Y to find $E[U^2]$. *Hint:* Your answer should depend on μ and σ.

(e) ______________

Solution:

$$E[U^2] = E[e^{2Y}] = E[e^{2t}] = m_Y(2) = e^{\mu t + \frac{\sigma^2 t^2}{2}}|_{t=2} = e^{2\mu + \frac{\sigma^2 2^2}{2}} = e^{2\mu + 2\sigma^2}$$

(f) (3 points) Find $V[U]$. *Hint:* Your answer should depend on μ and σ.

Solution:

$$V[U] = E[U^2] - E[U]^2 = e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2} = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

Solution: This is problem 6.111

6. (8 points) Fill in the blanks in the following statement of the central limit theorem. Note that *some variable names have changed* and you will have to adapt the statement accordingly.

   Let $Y_1, Y_2, \ldots, Y_n, \ldots$ be independent and identically distributed random variables with $E[Y_i] = a$ and $V[Y_i] = b^2$. Define

$$\bar{Y_n} = \underline{\frac{1}{n}(Y_1 + \cdots + Y_n)} \qquad \text{and} \qquad U_n = \frac{\bar{Y_n} - a}{\underline{(b/\sqrt{n})}}$$

   Then the distribution function of U_n converges to the <u>standard</u> normal distribution function as $n \to \infty$. That is,

$$\lim_{n\to\infty} P(\underline{U_n \le u}) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

   for all u.

7. One-hour carbon monoxide concentrations in air samples from a large city average 12 ppm (parts per million) with standard deviation 9 ppm.

   (a) (2 points) Are carbon monoxide concentrations in air samples from this city approximately normally distributed? Answer "yes" or "no".

   (a) ____________

   Solution: No.

   (b) (3 points) Why or why not?

   Solution: A normal distribution can take on any real value, and the probability that it is between 1.25 and 2 standard deviations below its mean is significant (around 10%). Here this is impossible, as carbon monoxide concentrations cannot be negative.

   (c) (2 points) Suppose we take 100 randomly selected air samples and look at the average carbon monoxide concentration in these samples. Is this average approximately normally distributed? Answer "yes" or "no".

   (c) ____________

   Solution: Yes.

(d) (3 points) Why or why not?

Solution: It is reasonable to assume that the samples are independent, and 100 is sufficiently large that we can apply the Central Limit Theorem.

(e) (7 points) Find the probability that the average concentration in 100 randomly selected samples will exceed 14 ppm.

(e) ____________

Solution: 0.0132.

Solution: This is problem 7.53

8. Scores on an examination are assumed to be normally distributed with mean 78 and variance 36.

(a) (6 points) Suppose that students scoring in the top 10% of this distribution are to receive an A grade. What is the minimum score a student must achieve to earn an A grade?

(a) ____________

Solution: We seek c such that $P(Y > c) = 0.1$. For the standard normal, $P(Z > z_0) = 0.1$ when $z_0 = 1.28$. So $c = 78 + (1.28)(6) = 85.68$.

(b) (8 points) If it is known that a student's score exceeds 72, what is the probability that his or her score exceeds 84?

(b) ____________

Solution:

$$P(Y > 84 \mid Y > 72) = \frac{P(Y > 84)}{P(Y > 72)} = \frac{P(Z > 1)}{P(Z > 1)} = \frac{0.1587}{0.8413} = 0.1886$$

Solution: This is problem 4.74

9. Let $Y_1, Y_2, \ldots, Y_5$ be a random sample of size 5 from a normal population with mean 0 and variance 1 and let

$$\bar{Y} = (1/5)\sum_{i=1}^{5} Y_i \qquad \text{and} \qquad W = \sum_{i=1}^{5} Y_i^2$$

(a) (4 points) What is the distribution of $\bar{Y}$ and what are the relevant parameters?

> **Solution:** $\bar{Y}$ is normal, with mean $\mu = 0$ and variance $\sigma^2 = 1/5$. You can either quote Theorem 7.1 or note that a linear combination of normal random variables is normal and then compute the means and variance yourself.

(b) (4 points) What is the distribution of W and what are the relevant parameters?

> **Solution:** W has the chi-squared distribution, with $\nu = 5$. This is an application of Theorem 7.2, (or 6.4).

(c) (4 points) What is the distribution of $\sum_{i=1}^{5}(Y_i - \bar{Y})^2$ and what are the relevant parameters?

> **Solution:** It is a chi-squared distribution, with $\nu = 4$. This is an application of Theorem 7.3.

> **Solution:** This is problem 7.37

.

10. Consider the following game: A player throws a fair die repeatedly until he rolls a 2, 3, 4, 5, or 6. In other words, the player continues to throw the die as long as he rolls 1s. When he rolls a "non-1," he stops. Let Y be the number of throws needed to obtain the first non-1.

(a) (4 points) Y is a random variable of a type we have studied. What is the name of this type of random variable and what are the relevant parameter(s)?

> **Solution:** Geometric with $p = 5/6$.

(b) (3 points) What is the probability that the player tosses the die exactly three times?

(b) ____________

> **Solution:** Compute the probability function for the geometric RV at $y = 3$ and get 0.023.

(c) (3 points) What is the expected number of rolls needed to obtain the first non-1?

(c) ____________

> **Solution:** The expectation of a geometric RV is $1/p = 1.2$.

(d) (8 points) If he rolls a non-1 on the first throw, the player is paid \$1. Otherwise, the payoff is doubled for each 1 that the player rolls before rolling a non-1. Thus, the player is paid \$2 if he rolls a 1 followed by a non-1; \$4 if he rolls two 1s followed by a non-1; \$8 if he rolls three 1s followed by a non-1; etc. In general, if the player rolls $(Y-1)$ 1s before rolling his first non-1, he is paid 2^{Y-1} dollars. What is the expected amount paid to the player?

(d) ______________

Solution: The expectation of 2^{Y-1} is

$$E[2^{Y-1}] = \sum_{y=1}^{\infty} 2^{y-1} p q^{y-1}$$

Sum the geometric series to get \$1.25.

Solution: This is problem 3.187

11. Suppose that two continuous random variables Y_1, Y_2 have joint density function given by

$$f(y_1, y_2) = \begin{cases} 6(1-y_2) & 0 \le y_1 \le y_2 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) (3 points) Graph the region of the plane in which f is nonzero.

(b) (6 points) Fill in the blanks to write down a double integral for $P(Y_1 \le 3/4)$. Note that the order of integration is prescribed.

Solution:

$$\int_0^{3/4} \int_{y_1}^{1} \underline{6(1-y_2)}\, dy_2 dy_1$$

(c) (12 points) Find $P(Y_2 \le 1/2 \mid Y_1 \le 3/4)$.

(c) ______________

Solution:

$$\frac{\int_0^{1/2} \int_{y_1}^{1/2} 6(1-y_2)\, dy_2 dy_1}{\int_0^{3/4} \int_{y_1}^{1} 6(1-y_2)\, dy_2 dy_1} = \frac{32}{63} \approx 0.5079$$

Solution: This is based on problem 5.27.

12. Suppose that Y_1, Y_2, Y_3 are three independent uniform random variables on the interval $[0, a]$. Let $F(y)$ be their common CDF and $f(y)$ their common pdf. Let $Y_{(3)} = \max(Y_1, Y_2, Y_3)$, and G, g be the CDF and PDF of Y_3.

(a) (6 points) Find $G(y)$ in terms of $F(y)$.

(a) ____________

Solution: $G(y) = P(Y_{(3)} \leq y) = P(Y_1 \leq y, Y_2 \leq y, Y_3 \leq y) = F(y)^3$.

(b) (6 points) Find $g(y)$ in terms of $F(y)$ and $f(y)$, using the previous part.

(b) ____________

Solution: Differentiate: $g(y) = 3F(y)^2 f(y)$.

(c) (6 points) Suppose that the number of minutes you have to wait for a bus is uniformly distributed on the interval $[0, 15]$. You wait for the bus 3 times. Write an integral for the probability that the longest wait is less than 10 minutes.

Solution: We can take $F(y) = y/15$, where $y \in [0, 15]$ and $f(y) = 1/15$ in the same interval. We get

$$\int_0^{10} g(y)\,dy = \int_0^{10} 3F(y)^2 f(y)\,dy = \int_0^{10} \frac{3y^2}{15^3}\,dy$$

(d) (3 points) Find the probability discussed in the previous part.

Solution:

$$\int_0^{10} \frac{3y^2}{15^3}\,dy = \frac{10^3}{15^3} \approx 0.2963$$

Solution: This is based on problem 6.76.

13. Suppose that X and Y are two independent exponential random variables with mean 1.

(a) (6 points) Graph the region in the (x, y) plane where the joint density function of X and Y is not zero and $y \leq 3x$.

(b) (10 points) Fill in the blanks to write down a double integral for $P(Y \leq 3X)$. Note that the order of integration is prescribed.

$$\int_0^{\infty} \int_{y/3}^{\infty} \underline{e^{-x}e^{-y}}\,dxdy$$

(c) (10 points) Find $P(Y \leq 3X)$.

Solution: Compute to get 3/4.