Math C067 Practice Questions

Richard Beigel

March 20, 2006

Note: you will be expected to show your work on the test. If you use a calculator, write down the formulas that you entered into the calculator. I have given some answers so that you can check your work. These answers would not be acceptable on a test without additional explanation.

Exercises:

1. (a) A 3-child family is chosen at random. What is the probability that a majority of the 3 children are girls? Assume that male and female children are equally likely and independent.

   Solution: The equiprobable sample space is

   $$\{BBB, BBG, BGB, BGG, GBB, GBG, GGB, GGG\}.$$

   Four of its points have more girls than boys. So the probability that a majority of the 3 children are girls is $\frac{4}{8} = \frac{1}{2}$.
   This can also be seen by symmetry. Boys and girls are equally likely. So a majority of girls has the same probability as a majority of boys. Letting A denote the event that a majority of the 3 children are girls, we have $p(A) = p(\overline{A})$ and $p(A)+p(\overline{A}) = 1$, so $p(A) = \frac{1}{2}$.

   (b) A 3-child family is chosen at random. We learn that at least one of the children in the family is a girl. What is the probability that a majority of the 3 children are girls? Assume that male and female children are equally likely and independent.

   Solution: The condition that at least one of the children is a girl reduces the sample space to $\{BBG, BGB, BGG, GBB, GBG, GGB, GGG\}$. Four of its points have more girls than boys. So the probability that a majoirty of the 3 children are girls is $\frac{4}{7}$.

   (c) A 3-child family is chosen at random. We learn that the oldest child is a girl. What is the probability that a majority of the 3 children are girls? Assume that male and female children are equally likely and independent.

Solution 1: In each triple, assume that the oldest child comes first. The condition that the oldest child is a girl reduces the sample space to $\{GBB, GBG, GGB, GGG\}$. Three of its points have more girls than boys. So the probability that a majoirty of the 3 children are girls is $\frac{3}{4}$.

Solution 2: We are told that the oldest child is a girl. So only the last two children are random. A majority of the 3 children are girls if and only if at least one of the last two is a girl. The sample space for the last two children is $\{BB, BG, GB, GG\}$. Three of those four points contain at least one girl, i.e., the probability is $\frac{3}{4}$.

2. Jack and Jill go for AIDS testing. Assume that Jack is from a high-risk group, i.e., the probability he has AIDS is 0.1. Assume that Jill is not from a high-risk group, i.e., the probability that she has AIDS is 0.001. Assume that Jack and Jill have been practicing safe sex with each other so far, i.e., Jack's AIDS status is independent of Jill's AIDS status. Assume that the test is accurate.

   (a) Scenario 1: While waiting for the results they get their ID numbers mixed up so they don't know whose result is whose. If exactly one of them tests positive, what is the probability that Jack is the one who tested positive?

   Solution: We define three events:

   - A is the event that Jack is positive
   - B is the event that Jill is positive
   - C is the event that exactly one of them is positive.

   We need to find $p(A|C)$. We will use Bayes' Theorem:

   $$p(A|C) \cdot p(C) = p(C|A) \cdot P(A).$$

   We know

   - $p(A) = 0.1$
   - $p(B) = 0.001$

   We still need to find $p(C)$ and $p(C|A)$.

   $$\begin{aligned} C &= (A \cap \overline{B}) \cup (\overline{A} \cap B) \\ p(C) &= p(A \cap \overline{B}) + p(\overline{A} \cap B) && \text{disjoint union} \\ p(C) &= p(A) \cdot p(\overline{B}) + p(\overline{A}) \cdot p(B) && \text{independence} \\ p(C) &= p(A) \cdot (1 - p(B)) + (1 - p(A)) \cdot p(B) \\ p(C) &= 0.1 \cdot 0.999 + 0.9 \cdot 0.001 \\ p(C) &= 0.1008 \end{aligned}$$

We find $p(C|A)$ as follows:

$$\begin{aligned}
C &= (A \cap \overline{B}) \cup (\overline{A} \cap B) \\
p(C|A) &= p(A \cap \overline{B}|A) + p(\overline{A} \cap B|A) \qquad \text{disjoint union} \\
p(C|A) &= p(\overline{B}|A) + 0 \\
p(C|A) &= p(\overline{B}|A) \qquad \text{you can also see this directly via logic} \\
p(C|A) &= p(\overline{B}) \\
p(C|A) &= 1 - p(B) \\
p(C|A) &= 0.999
\end{aligned}$$

Finally,

$$\begin{aligned}
p(A|C) \cdot p(C) &= p(C|A) \cdot P(A) \\
p(A|C) \cdot 0.1008 &= 0.999 \cdot 0.1 \\
p(A|C) &= 0.999 \cdot 0.1/0.1008 \\
p(A|C) &\approx 0.991071429
\end{aligned}$$

(b) (Still Scenario 1). What is the probability that Jill is the one who is sick?

Solution: Since the condition C says that exactly one of them is sick, $p(B|C) = 1 - p(A|C) = 0.008928571$.

(c) Scenario 2: To save money Jack and Jill's blood samples are mixed together and tested. Assume that the combined sample tests positive, i.e., at least one of them is sick. What is the probability that Jack is sick?

Solution:

- A and B are defined as above.
- C is the event that at least one of them is positive.

We need to find $p(A|C)$. We will use Bayes' Theorem:

$$p(A|C) \cdot p(C) = p(C|A) \cdot P(A).$$

We know

- $p(A) = 0.1$
- $p(B) = 0.001$

We still need to find $p(C)$ and $p(C|A)$.

$$\begin{aligned}
C &= A \cup B \\
p(C) &= p(A) + p(B) - p(A \cap B) \\
p(C) &= p(A) + p(B) - p(A \cap B) \\
p(C) &= p(A) + p(B) - p(A) \cdot p(B) \qquad \text{independence} \\
p(C) &= 0.1 + 0.001 - 0.0001 \\
p(C) &= 0.1009
\end{aligned}$$

$p(C|A) = 1$ because if Jack is positive then at least one of them is positive. Finally,

$$\begin{aligned} p(A|C) \cdot p(C) &= p(C|A) \cdot P(A) \\ p(A|C) \cdot 0.1009 &= 1 \cdot 0.1 \\ p(A|C) &= 1 \cdot 0.1/0.1009 \\ p(A|C) &\approx 0.991080277 \end{aligned}$$

(d) (Still Scenario 2). What is the probability that Jill is sick?
By the same reasoning as above, $p(C|B) = 1$ so

$$\begin{aligned} p(B|C) \cdot p(C) &= p(C|B) \cdot P(B) \\ p(B|C) \cdot 0.1009 &= 1 \cdot 0.001 \\ p(B|C) &= 1 \cdot 0.001/0.1009 \\ p(B|C) &\approx 0.00991080277 \end{aligned}$$

(e) (Still Scenario 2). What is the probability that they are both sick? By the same reasoning as above, $p(C|B \cap A) = 1$ so

$$\begin{aligned} p(B \cap A|C) \cdot p(C) &= p(C|B \cap A) \cdot P(B \cap A) \\ p(B \cap A|C) \cdot 0.1009 &= 1 \cdot P(B) \cdot P(A) \qquad \text{independence} \\ p(B \cap A|C) \cdot 0.1009 &= 1 \cdot 0.0001 \\ p(B \cap A|C) &= 1 \cdot 0.0001/0.1009 \\ p(B \cap A|C) &\approx 0.000991080277 \end{aligned}$$

3. A fair coin is tossed until a head or three tails appears. Let X be a random variable that denotes the number of coin tosses made.

(a) What is the range of X?

(b) What is the distribution of X?

(c) Find $E[X]$.

(d) Find $\mathrm{Var}[X]$.

(e) Find σ_X.

Solution.

(a) The set of possible values for X is $\{1, 2, 3\}$

(b) What is the distribution of X?

- $p(X = 1) = p(H) = \frac{1}{2}$
- $p(X = 2) = p(TH) = \frac{1}{4}$
- $p(X = 3) = p(TT) = \frac{1}{4}$

x	1	2	3
$f(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

(c)

$$E[X] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} = \frac{7}{4}.$$

(d)

$$\mathrm{Var}[X] = (1 - \frac{7}{4})^2 \cdot \frac{1}{2} + (2 - \frac{7}{4})^2 \cdot \frac{1}{4} + (3 - \frac{7}{4})^2 \cdot \frac{1}{4} = 0.6875.$$

(e)

$$\sigma_X = \sqrt{0.6875} \approx 0.829156198$$

4. A biased coin is spun 3 times. The probability that it comes up heads is $\frac{1}{3}$. What is the probability that we get

   (a) exactly one heads?
   (b) at least one heads?
   (c) three heads?

5. A biased coin is spun 2500 times. The probability that it comes up heads is $\frac{1}{5}$. Let X denote the number of times the coin comes up heads. Let Y denote the number of times the coin comes up tails. What is

   (a) the expected value of X?
   (b) the variance of X?
   (c) the standard deviation of X?
   (d) the expected value of Y?
   (e) the variance of Y?
   (f) the standard deviation of Y?

   Solution: In parts a–c we work with $B(2500, \frac{1}{5})$. In parts d–f we work with $B(2500, \frac{4}{5})$.

   (a) $E[X] = 2500 \cdot \frac{1}{5} = 500$.
   (b) $\mathrm{Var}[X] = 2500 \cdot \frac{1}{5}\frac{4}{5} = 400$.
   (c) $\sigma_X = \sqrt{\mathrm{Var}[X]} = \sqrt{400} = 20$.
   (d) $E[Y] = 2500 \cdot \frac{4}{5} = 2000$.
   (e) $\mathrm{Var}[Y] = 2500 \cdot \frac{4}{5}\frac{1}{5} = 400$.

(f) $\sigma_Y = \sqrt{\mathrm{Var}[Y]} = \sqrt{400} = 20$.

6. A biased coin is spun 2500 times. The probability that it comes up heads is $\frac{1}{5}$. Let X denote the number of times the coin comes up heads. Estimate:

   (a) $P(X = 500)$

   (b) $P(200 \leq X \leq 400)$

   (c) $P(300 \leq X \leq 500)$

   (d) $P(400 \leq X \leq 600)$

   (e) $P(500 \leq X \leq 700)$

   (f) $P(600 \leq X \leq 800)$

   (g) $P(X \geq 400)$

   (h) $P(X \geq 500)$

   (i) $P(X \geq 600)$

   (j) $P(X \leq 400)$

   (k) $P(X \leq 500)$

   (l) $P(X \leq 600)$

7. We work with $B(2500, \frac{1}{5})$.

   (a)

$$\begin{aligned} P(X = 500) &= P(499.5 \leq X \leq 500.5) \\ &= P(499.5 \leq X \leq 500) + P(500 < X \leq 500.5) \end{aligned}$$

   The first interval consists of $\frac{500-499.5}{20} = 0.025$ standard deviations. The second interval consists of $\frac{500.5-500}{20} = 0.025$ standard deviations. 0.025 isn't in our table so we average the entries for 0.02 and 0.03 in Table A-1 to get $(0.0080 + 0.0120)/2 = 0.0100$. (In class we used a table with more digits of accurancy.) Therefore $P(X = 500) \approx 0.0100 + 0.0100 = 0.0200$.

   (b)

$$\begin{aligned} P(200 \leq X \leq 400) &= P(199.5 \leq X \leq 400.5) \\ &= P(199.5 \leq X \leq 500) - P(400.5 < X \leq 500) \end{aligned}$$

   The first interval consists of $\frac{500-199.5}{20} = 15.025$ standard deviations. The second interval consists of $\frac{500-400.5}{20} = 4.975$ standard deviations. Table A-1 stops at 3.99 because beyond that the probability is 0.5000 to four decimal places. Therefore $P(200 \leq X \leq 400) \approx 0.5000 - 0.5000 = 0$.

   (c)

$$\begin{aligned} P(300 \leq X \leq 500) &= P(299.5 \leq X \leq 500.5) \\ &= P(299.5 \leq X \leq 500) + P(500 < X \leq 500.5) \end{aligned}$$

The first interval consists of $\frac{500-299.5}{20} = 10.025$ standard deviations. The second interval consists of $\frac{500.5-500}{20} = 0.025$ standard deviations. Therefore $P(300 \leq X \leq 500) \approx 0.5000 + 0.0100 = 0.5100$.

(d)

$$\begin{aligned} P(400 \leq X \leq 600) &= P(399.5 \leq X \leq 600.5) \\ &= P(399.5 \leq X \leq 500) + P(500 < X \leq 600.5) \end{aligned}$$

The first interval consists of $\frac{500-399.5}{20} = 5.025$ standard deviations. The second interval consists of $\frac{600.5-500}{20} = 5.025$ standard deviations. Therefore $P(400 \leq X \leq 600) \approx 0.5000 + 0.5000 = 1.000$.

(e)

$$\begin{aligned} P(500 \leq X \leq 700) &= P(499.5 \leq X \leq 700.5) \\ &= P(499.5 \leq X \leq 500) + P(500 < X \leq 700.5) \end{aligned}$$

The first interval consists of $\frac{500-499.5}{20} = 0.025$ standard deviations. The second interval consists of $\frac{700.5-500}{20} = 10.025$ standard deviations. Therefore $P(500 \leq X \leq 700) \approx 0.0100 + 0.5000 = 0.5100$.

(f) $P(600 \leq X \leq 800)$

$$\begin{aligned} P(600 \leq X \leq 800) &= P(599.5 \leq X \leq 800.5) \\ &= P(500 \leq X \leq 800.5) - P(500 \leq X < 599.5) \end{aligned}$$

The first interval consists of $\frac{500-199.5}{20} = 15.025$ standard deviations. The second interval consists of $\frac{500-400.5}{20} = 4.975$ standard deviations. Therefore $P(600 \leq X \leq 800) \approx 0.5000 - 0.5000 = 0$.

(g) $P(X \geq 400)$

$$\begin{aligned} P(X \geq 400 &= P(X \geq 399.5) \\ &= P(399.5 \leq X \leq 500) + P(X > 500) \end{aligned}$$

The first interval consists of $\frac{500-399.5}{20} = 5.025$ standard deviations. The second interval consists of half of the distribution (infinitely many standard deviations). Therefore $P(X \geq 400) \approx 0.5000 + 0.5000 = 1.000$.

(h) $P(X \geq 500)$

$$\begin{aligned} P(X \geq 500) &= P(499.5 \leq X) \\ &= P(499.5 \leq X \leq 500) + P(500 < X) \end{aligned}$$

The first interval consists of $\frac{500-499.5}{20} = 0.025$ standard deviations. The second interval consists of half of the distribution. Therefore $P(X \geq 500) \approx 0.0100 + 0.5000 = 0.5100$.

(i) $P(X \geq 600)$

$$\begin{aligned} P(X \geq 600) &= P(X \geq 599.5) \\ &= P(500 \leq X) - P(500 \leq X < 599.5) \end{aligned}$$

The first interval consists of half of the distribution. The second interval consists of $\frac{599.5-500}{20} = 4.975$ standard deviations. Therefore $P(X \geq 600) \approx 0.5000 - 0.5000 = 0$.

(j)

$$\begin{aligned} P(X \leq 400) &= P(X \leq 400.5) \\ &= P(X \leq 500) - P(400.5 < X \leq 500) \end{aligned}$$

The first interval consists of half of the distribution. The second interval consists of $\frac{500-400.5}{20} = 4.975$ standard deviations. Therefore $P(X \leq 400) \approx 0.5000 - 0.5000 = 0$.

(k) $P(X \leq 500)$

$$\begin{aligned} P(X \leq 500) &= P(X \leq 500.5) \\ &= P(X \leq 500) + P(500 < X \leq 500.5) \end{aligned}$$

The first interval consists of half of the distribution. The second interval consists of $\frac{500.5-500}{20} = 0.025$ standard deviations. Therefore $P(X \leq 500) \approx 0.5000 + 0.0100 = 0.5100$.

(l) $P(X \leq 600)$

$$\begin{aligned} P(X \leq 600) &= P(X \leq 600.5) \\ &= P(X \leq 500) + P(500 < X \leq 600.5) \end{aligned}$$

The first interval consists of half of the distribution. The second interval consists of $\frac{600.5-500}{20} = 5.025$ standard deviations. Therefore $P(X \leq 600) \approx 0.5000 + 0.5000 = 1.000$.

8. A die is rolled 45 times. Let X denote the number of times that it comes up 5. What is

(a) the expected value of X?

(b) the variance of X?

(c) the standard deviation of X?

Solution: The distribution of X is $B(45, \frac{1}{6})$.

(a) $E[X] = 45 \cdot \frac{1}{6} = 7.5$.

(b) $\mathrm{Var}[X] = 45 \cdot \frac{1}{6}\frac{5}{6} = 6.25$.

(c) $\sigma_X = \sqrt{\mathrm{Var}[X]} = \sqrt{6.25} = 2.5$.

9. A die is rolled 45 times. Let X denote the number of times that it comes up 3. Estimate

   (a) $P(X = 7.5)$
   (b) $P(5 \leq X \leq 10)$
   (c) $P(10 \leq X \leq 20)$
   (d) $P(X \leq 5)$
   (e) $P(X \leq 8)$
   (f) $P(X \leq 15)$
   (g) $P(X \geq 5)$
   (h) $P(X \geq 8)$
   (i) $P(X \geq 15)$

   Solution: The distribution of X is $B(45, \frac{1}{6})$.

   (a) 7.5 is not in the range of X, so $P(X = 7.5) = 0$.

   (b)
$$\begin{aligned} P(5 \leq X \leq 10) &= P(4.5 \leq X \leq 10.5) \\ &= P(4.5 \leq X \leq 7.5) + P(7.5 < X \leq 10.5) \end{aligned}$$
   The first interval consists of $\frac{7.5-4.5}{2.5} = 1.2$ standard deviations. The second interval consists of $\frac{10.5-7.5}{2.5} = 1.2$ standard deviations. Therefore, $P(5 \leq X \leq 10) \approx 0.3849 + 0.3849 = 0.7698$.

   (c)
$$\begin{aligned} P(10 \leq X \leq 20) &= P(9.5 \leq X \leq 20.5) \\ &= P(7.5 \leq X \leq 20.5) - P(7.5 \leq X < 9.5) \end{aligned}$$
   The first interval consists of $\frac{20.5-7.5}{2.5} = 5.2$ standard deviations. The second interval consists of $\frac{9.5-7.5}{2.5} = 0.8$ standard deviations. Therefore, $P(5 \leq X \leq 10) \approx 0.5000 - 0.2881 = 0.2119$.

   (d)
$$\begin{aligned} P(X \leq 5) &= P(X \leq 5.5) \\ &= P(X \leq 7.5) - P(5.5 < X \leq 7.5) \end{aligned}$$
   The first interval consists of half of the distribution. The second interval consists of $\frac{7.5-5.5}{2.5} = 0.8$ standard deviations. Therefore, $P(X \leq 5) \approx 0.5000 - 0.2881 = 0.2119$.

(e)

$$\begin{aligned} P(X \leq 8) &= P(X \leq 8.5) \\ &= P(X \leq 7.5) + P(7.5 < X \leq 8.5) \end{aligned}$$

The first interval consists of half of the distribution. The second interval consists of $\frac{8.5-7.5}{2.5} = 0.4$ standard deviations. Therefore, $P(X \leq 8) \approx 0.5000 + 0.1554 = 0.6554$.

(f)

$$\begin{aligned} P(X \leq 15) &= P(X \leq 15.5) \\ &= P(X \leq 7.5) + P(7.5 < X \leq 15.5) \end{aligned}$$

The first interval consists of half of the distribution. The second interval consists of $\frac{15.5-7.5}{2.5} = 3.2$ standard deviations. Therefore, $P(X \leq 15) \approx 0.5000 + 0.4993 = 0.9993$.

(g)

$$\begin{aligned} P(X \geq 5) &= P(X \geq 4.5) \\ &= P(4.5 \leq X \leq 7.5) + P(7.5 \leq X) \end{aligned}$$

The first interval consists of $\frac{7.5-4.5}{2.5} = 1.2$ standard deviations. The second interval consists of half of the distribution. Therefore, $P(X \geq 5) \approx 0.5000 + 0.3849 = 0.8849$.

(h)

$$P(X \geq 8) = P(X \geq 7.5)$$

That interval consists of half of the distribution. Therefore, $P(X \geq 8) \approx 0.5000$.

(i) $P(X \geq 15)$

$$\begin{aligned} P(X \geq 15) &= P(X \geq 14.5) \\ &= P(7.5 \leq X) - P(7.5 \leq X < 14.5) \end{aligned}$$

The first interval consists of half of the distribution. The second interval consists of $\frac{14.5-7.5}{2.5} = 2.8$ standard deviations.
Therefore, $P(X \geq 15) \approx 0.5000 - 0.4974 = 0.0026$.

10. A roulette wheel has slots numbered 0, 00, and 1 through 36. Assume that the wheel is fair so that each number has the same probability of coming up.

(a) On a single spin, what is the probability of getting 17?

Answer: Approximately 0.0263157895

(b) On a single spin, what is the probability of not getting 17?

Answer: Approximately 0.973684211

(c) The wheel is spun 38 times. What is the probability of not getting any 17s?

Answer: Approximately 0.362985137

(d) The wheel is spun 38 times. What is the probability of getting at least one 17?

Answer: Approximately 0.637014863

(e) The wheel is spun 38 times. What is the expected number of 17s?

Answer: 1

(f) Consider the experiment of spinning the wheel 38 times and counting how many 17s you get. What is the variance in that experiment?

Answer: 0.973684211

(g) What is the standard deviation?

Answer: 0.986754382

(h) **Extra Credit** The wheel is spun 38 times. Estimate the probability of getting at least two 17s.

Solution: The mean $np = 1$, which is too small for us to approximate the binomial distribution by a normal distribution. When np is small and n is large, the binomial distribution is approximated by the Poisson distribution (pp. 190–191) with $\lambda = np$. We have

$$P(X = k) \approx \lambda^k e^{-\lambda}/k! = e^{-1}/k!$$

- $p(X = 0) \approx e^{-1}/0! = e^{-1} \approx 0.367879441$
- $p(X = 1) \approx e^{-1}/1! = e^{-1} \approx 0.367879441$
- $p(X \leq 1) = p(X = 0) + p(X = 1) \approx 0.367879441 + 0.367879441 = 0.735758882$
- $p(X \geq 2) = 1 - p(X \leq 1) \approx 1 - 0.735758882 = 0.264241118$

(i) The wheel is spun 38 times. We learn that there was at least one 17. Given that information, estimate the probability that there were at least two 17s. Note: you will need an answer for the preceding part in order to do this part. If you didn't answer the preceding part, **guess an answer** for it. You can still get credit for this part even if you got the preceding part wrong.

Solution:

$$
\begin{aligned}
p(X \geq 2|X \geq 1) &= p(X \geq 2 \text{ and } X \geq 1)/P(X \geq 1) \\
&= p(X \geq 2)/P(X \geq 1) \\
&\approx 0.264241118/0.637014863 \\
&\approx 0.414811543
\end{aligned}
$$