401 Midterm 2 Solution

1. (20 pts) Answer T (true) or F (false) to the following statements.

- It is always the case that $P(A|B) = P(B|A)$. False.
- You can talk of conditional probability $P(A|B)$ for any set B. False. If $P(B) = 0$ we cannot discuss $P(A|B)$.
- $P(A|B) = 1$ if B is a subset of A. True. Because $P(A \cap B) = P(B)$.
- $c + X$ and X has the same variance for any c. True. Shifting does not change variance.
- If A is independent of B and B is independent of C, A is independent of C. False. You can think of an example where $A = C$.
- If two events are disjoint, they cannot be independent. False. $P(\emptyset \cap \Omega) = 0 = P(\emptyset)P(\Omega)$.
- An event A can be independent of itself. True. $P(\emptyset \cap \emptyset) = 0 = P(\emptyset)P(\emptyset)$.
- The pmf is only defined for random variables that take a finite set of possible values. False. It can take infinite set of possible values as long as it is countable.
- Even if you change the value of the pdf on a countably infinite number of points, you still get the same distribution. True. Integration ignores values from a countable number of points.
- Uniform distributions are closed under linear transformation. That is, if we have U, a uniform from a to b, $cU + d$ is also going to be a uniform. True. $cU + d \sim U(d + ca, d + cb)$. If c is negative, switch the two ends.

2. (20 pts) For each of the following problems, choose the most suitable option.

- If you take a uniform from 0 to 1, and map it with a negative of a logarithm, you get back a,

  (a) geometric distribution
  (b) exponential distribution
  (c) normal distribution
  (d) binomial distribution

  The answer is b. This was one of the homework problems. Write the cdf and apply a function to both sides of the inequality.

- If a continuous distribution takes values on a region from 0 to 0.5, there must be some x between 0 and 0.5 such that,

  (a) $f(x) \geq 2$
  (b) $f(x) = 0$
  (c) $f(x) < 2$
  (d) $f(x) < 0$

The answer is a. If a was not true, then the integral of the pdf on the range will be smaller than 1.

- Generate points within a sphere with radius 1 uniformly. Record the distance of the point from the origin. The range of this random variable is,

  (a) $\{x, y, z : x^2 + y^2 + z^2 \leq 1\}$
  (b) $[-1, 1]$
  (c) $[0, 1]$
  (d) $\{x, y : x^2 + y^2 \leq 1\}$

  The answer is c. Because it is within the sphere of radius 1, the distance cannot be larger than 1. Also distance has to be nonnegative.

- In the above setting, probability of getting a point with radius smaller than $0 < k < 1$ is,

  (a) k over 1
  (b) area of disc with radius k over area of disc with radius 1
  (c) volume of sphere with radius k over volume of sphere with radius 1
  (d) none of the above

  The answer is c. The wording is not ideal. I meant probability of getting a point with distance from the origin smaller than k. Just think about the regions in the sphere that give a distance smaller than k.

- You throw two die. Record the larger of the two as the outcome of this random variable. What is its expected value?

  (a) $-\frac{5}{2}$
  (b) $\frac{20}{3}$
  (c) $\frac{160}{35}$
  (d) $\frac{161}{36}$

  Please find the solution as an article on my website.

3. (10 pts) Explain how to reproduce a fair die from repeated flips of a fair coin. That is you take a coin, repeatedly flip them, and assign numbers 1 to 6 to the outcomes so that the probability of getting each number is the same. There is no restriction on how many times you flip the coin. You may be able to assign a number with just 3 flips or it might be that it takes much more flips like 1000. Flip a coin three times. Code H as 1, T as 0. Read the result as a binary. If you get a 0 or a 7, discard the result and flip another set of 3 coins.

4. (10 pts) Explain how to reproduce a fair random generator of numbers from 1 to 10 with repeated throws of a fair die. That is you take a die, repeatedly throw them, and assign numbers 1 to 10 to the outcomes so that probability of getting each number is the same. There is no restriction on how many times you throw the die. You may be able to assign a number with just 3 throws or it might be that it takes much more throws like 1000. Throw a die once and see if it is 1, 2, 3 or 4, 5, 6. If it is the former read the subsequent dice throws as is. If it is the latter read the subsequent dice throws with 5 added to it. If you throw a 6 reroll.

5. (10 pts) What is the expected number of coin flips needed to reproduce a dice throw outcome in problem 3? This is geometric with success probability $\frac{3}{4}$ for each set of 3 flips. So we expect $\frac{4}{3}$ sets of flips of 3 before we get an outcome. Multiplying 3 to this we get 4.

6. (10 pts) What is the expected number of dice throws needed to reproduce a random number from 1 to 10 in problem 4? The first throw we will do for sure. The subsequent throws are geometric with success probability $\frac{5}{6}$. So we expect $\frac{6}{5}$ throws in the subsequent throws. In total it is $\frac{11}{5}$.

7. (5 pts) Think of the first eight prime numbers. A random variable takes these numbers with probability proportional to its size. Calculate its pmf. The first 8 primes are 2, 3, 5, 7, 11, 13, 17, 19. Add them all up and we get 77. So we have $p(2) = \frac{2}{77}$, $p(3) = \frac{3}{77}$, $p(5) = \frac{5}{77}$, $p(7) = \frac{7}{77} = \frac{1}{11}$, $p(11) = \frac{11}{77} = \frac{1}{7}$, $p(13) = \frac{13}{77}$, $p(17) = \frac{17}{77}$, $p(19) = \frac{19}{77}$.

8. (5 pts) Calculate the variance of the above random variable. $E[X] = \frac{4+9+25+49+121+169+289+361}{77} = \frac{13+25+49+290+650}{77} = \frac{62+25+940}{77} = \frac{1027}{77}$. $E[X^2] = \frac{8+27+125+343+1331+2197+4913+6859}{77} = \frac{15803}{77}$. $\frac{15803}{77} - \left(\frac{1027}{77}\right)^2 =$ 27.34.

9. (5 pts) The pdf of a random variable supported from 1 to ∞ is $k\frac{1}{x^4}$. Where k is some constant. Find k. $k\int_1^\infty x^{-4}dx = 1$. so $k\left[-\frac{1}{3}x^{-3}\right]_1^\infty = 1$. Therefore $k = 3$.

10. (5 pts) Find the variance of the random variable in the above problem. $E[X] = 3\int_1^\infty x^{-3}dx = 3\left[-\frac{1}{2}x^{-2}\right]_1^\infty = \frac{3}{2}$. $E[X^2] = 3\int_1^\infty x^{-2}dx = 3\left[-x^{-1}\right]_1^\infty = 3$. Therefore $3 - \frac{9}{4} = \frac{3}{4}$.