SDS 321: Practice midterm

Note: This practice exam is longer than your actual exam will be! The actual exam will be 5 questions

1. Find a formula for the probability that among a set of n people, **at least two have their birthdays in the same month** of the year (assuming the months are equally likely for birthdays). Assume $n \leq 12$ (for $n > 13$ the probability is equal to 1.0). Let A be the event "at least one matching b-day month". It is easier to find $\mathbf{P}(A^c)$ – the probability of"no matching b-day months." First count the number of outcomes in A^c. It is $12 \cdot 11 \ldots (12-(n-1)) = 12!/(12-n)!$. The number of possible ordered combinations of birthday monts is 12^n. Therefore $\mathbf{P}(A^c) = \frac{12!/(12-n)!}{12^n}$ and $p(A) = 1 - p(A) = 1 - \frac{12!/(12-n)!}{12^n}$

2. How many ways can we arrange the letters of the word "ALPHABET" so that the first and last letters are vowels? There are three vowels - A,A,E. So, 4 possible ways of having vowels at both ends: A******A, A******E, E******A. For each of these combinations, we have 6 remaining, different, letters... so in each case, ****** can be any of 6! permutations of the remaining letters. So, there are $3 \cdot 6! = 2160$ combinations.

3. If a coin is tossed a sequence of times (infinitely many times), what is the probability that the first head will occur **after** the 5th toss, given that it has not occurred in the first 2 tosses? We know the first 2 coin tosses are tails. The probability of this is $\mathbf{P}(T,T) = (1-p)^2$ We want the probability that the first 5 are tails, given the first two are heads. The unconditional probability $\mathbf{P}(T,T,T,T,T) = (1-p)^5$. So, $\mathbf{P}((T,T,T,T,T)|(T,T)) = \frac{\mathbf{P}(T,T,T,T,T)}{\mathbf{P}(T,T)} = (1-p)^3$

4. A continuous random variable has PDF

$$f_X(x) = \begin{cases} a + bx^2 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

   (a) If $E[X] = \frac{3}{5}$, find a and b. PDF must integrate to 1, so $\int_0^1 a + bx^2 dx = \left[ax + \frac{b}{3}x^3\right]_0^1 = a + b/3 = 1$, so $b = 3(1-a)$
   $E[X] = 3/5$, so $\int_0^1 ax + bx^3 dx = \left[\frac{a}{2}x^2 + \frac{b}{4}x^4\right]_0^1 = \frac{a}{2} + \frac{b}{4} = \frac{3}{5}$

So,

$$\begin{aligned}\frac{a}{2}+\frac{3(1-a)}{4}&=\frac{3}{5}\\ 10a+15(1-a)&=12\\ a&=\frac{3}{5}\\ b=3(1-a)&=\frac{6}{5}\end{aligned}$$

(b) What is $\text{var}(X)$? If you couldn't find a and b, you can leave them as algabraic variables. $\text{var}(X)=E[X^2]-E[X]^2$. $E[X^2]=\int_0^1 x^2 f_X(x)dx=\int_0^1 ax^2+bx^4dx=\left[\frac{ax^3}{3}+\frac{bx^5}{5}\right]_0^1=\frac{a}{3}+\frac{b}{5}=\frac{1}{5}+\frac{6}{25}=\frac{11}{25}$, so $\text{var}(X)=\frac{11}{25}-\left(\frac{3}{5}\right)^2=\frac{11-9}{25}=\frac{2}{25}$.

(c) If X and Y are continuous random variables, show that, if X and Y are not independent, $\text{var}(X+Y)=\text{var}(X)+\text{var}(Y)+2(E[XY]-E[X]E[Y])$. $\text{var}(X+Y)=E[(X+Y)^2]-E[X+Y]^2=E[X^2+2XY+Y^2]-(E[X]+E[Y])^2=E[X^2]+2E[XY]+E[Y^2]-E[X]^2-2E[X]E[Y]-E[Y]^2$ (by linearity of expectation) $=\text{var}(X)+\text{var}(Y)+2(E[XY]-E[X]E[Y])$.

5. X is a normal random variable with mean 20 and variance 9.

(a) Write $\mathbf{P}(X\leq 15)$ in terms of the standard normal distribution, and use the attached standard normal table to find $\mathbf{P}(X\leq 15)$ $\mathbf{P}(X\leq 15)=\mathbf{P}\left(Z\leq\frac{15-20}{3}\right)=\mathbf{P}(Z\leq -1.67)=1-\mathbf{P}(Z\leq 1.67)=1-0.9525=0.0425$

(b) Write $\mathbf{P}(15\leq X\leq 20)$ in terms of the standard normal distribution, and use the attached standard normal table to find $\mathbf{P}(15\leq X\leq 20)$. $\mathbf{P}(15\leq X\leq 20)=\mathbf{P}(-5/3\leq Z\leq 0)=\mathbf{P}(Z\leq 0)-\mathbf{P}(Z\leq -5/3)=0.5-0.0425=0.4575$

(c) If Y is a normal random variable with mean 2 and variance 12, what is the mean and variance of $Z=2X+Y+3$? $E[Z]=2E[X]+E[Y]+3=40+2+3=45$. X and Y are independent, so $\text{var}(Z)=4\text{var}(X)+\text{var}(Y)=4\cdot 9+12=48$.

6. On average, 3 traffics accidents occur in a given city per day. You may assume that the number of accidents follows a Poisson distribution with $\lambda = 3$. Recall that the PMF for the Poisson distribution is $p_X(k) = \frac{\lambda^k e^{-\lambda}}{k!}$.

   (a) What is the probability that we see at least three accidents in a day? $\mathbf{P}(X \geq 3) = 1 - \mathbf{P}(X = 0) - \mathbf{P}(X = 1) - \mathbf{P}(X = 2) = 1 - e^{-3}(1 + 3 + 3^2/2) = 0.577$

   (b) If you know there is at least one accident, what is the probability that the total number of accidents is at least three? $\mathbf{P}(X \geq 1) = 1 - \mathbf{P}(X = 0) = 1 - e^{-3} = 0.950$. $\mathbf{P}(X \geq 3 | X \geq 1) = \mathbf{P}(X \geq 3)/\mathbf{P}(X \geq 1) = 0.577/0.950 = 0.607$

7. A continuous r.v. X has density function

$$f(x) = c\, e^{|x|} = \begin{cases} c\, e^{-x} & x \geq 0 \\ c\, e^{x} & x < 0 \end{cases}$$

   (a) Find c. $\int_{\infty}^{\infty} f_X(x) = 1$. By symmetry, $0.5 = c\int_0^{\infty} e^{-x} = c[-e^{-x}]_0^{\infty} = c \rightarrow c = 0.5$.

   (b) Find $\mathbf{P}(|X| < 2)$. $\mathbf{P}(-2 < X < 2) = 2\int_0^2 f_X(x)dx = \frac{2}{2}\int_0^2 e^{-x}dx = 1 - e^{-2} = 0.86$

   (c) Find $E(X)$ (*Hint:* no integral required). By symmetry, $E[X] = 0$

8. Charles claims that he can distinguish between whisky and brandy 75% of the time. Let $p =$ Charles' probability of distinguishing the drinks. Charles' claim is $p = p_C = 0.75$.
   Ruth bets that he cannot tell the difference and, in fact, just guesses. That is, Ruth's claim is $p = p_R = 0.5$.
   To settle this, a bet is made:
   Charles is to be given $n = 5$ small glasses, each having been filled with whisky or brandy, chosen by tossing a fair coin. He wins the bet if he gets **4 or more** correct.

   (a) If Charles is correct—i.e. if $p = p_C = 0.75$—what is the probability that Charles wins the bet. The question is for $\mathbf{P}(X \geq 4)$ when $p = p_C$. In that case $X \sim \mathit{Binomial}(n, p_C)$ and $\mathbf{P}(X \geq 4) = p_X(4) + p_X(5) = 5p_C^4(1 - p_C) + p_C^5 = p_C^4(5 - 4p_C) = 0.63$

   (b) If Ruth is correct—i.e. if Charles is guessing randomly—what is the probability that Charles wins the bet.

   The question is for $\mathbf{P}(X \geq 4)$ when $p = p_R$. Now $X \sim \mathit{Binomial}(n, p_R)$ and $\mathbf{P}(X \geq 4) = p_X(4) + p_X(5) = 5p_R^4(1 - p_R) + p_R^5 = 6\left(\frac{1}{2}\right)^5 = 0.19$

   (c) Assume that, before seeing the outcome of the bet, you believe that Charles is right with probability 0.1, and Ruth is right with probability 0.9

   I find out that Charles won the bet. With what probability should I believe Charles is right?

   Let A be "Charles is correct" and B be "Charles wins the bet". $P(A) = 0.1$, $P(B|A) = 0.63$, $P(B|A^c) = 0.19$, so

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(B|A)\mathbf{P}(A)}{\mathbf{P}(B|A)\mathbf{P}(A) + \mathbf{P}(B|A^c)\mathbf{P}(A^c)} = \frac{.1 \cdot .63}{.1 \cdot .63 + .9 \cdot .19} = 0.27$$

Standard normal table

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990