Math 10B with Professor Stankova
Worksheet, Midterm #2; Wednesday, 3/21/2018
GSI name: Roy Zhao

1 Problems

1.1 Bayes Theorem

1. Suppose a test is 99% accurate and 1% of people have a disease. What is the probability that you have the disease given that you tested positive?

2. Suppose I have 3 boxes with 10 red and blue balls in each but with different distributions. Box 1 has 2 red and 8 blue, box 2 has 5 red and 5 blue, and box 3 has 8 red and 2 blue. I randomly pick a box without looking and then pick out a ball. What is the probability that I selected box 3 given that I drew out a red ball?

3. Suppose I have 3 die. One is 4 sided with numbers $1-4$, one is 6 sided with numbers $1-6$, and one is 10 sided with sides $1-10$. I randomly pick a die without looking and then roll it. What is the probability that I selected the 4 sided die given that I rolled a 1?

1.2 Random Variables

4. Suppose that we roll two die and let X be equal to the maximum of the two rolls. Find $P(X \in \{1,3,5\})$ and draw the PMF for X.

5. I draw 5 cards from a deck of cards. Let X be the number of hearts I draw. What is the range of X and draw the PMF of X. Use this to find the probability that I draw at least 2 hearts.

1.3 Discrete Distributions

6. I am picking cards out of a deck. What is the probability that I pull out 2 kings out of 8 cards if I pull with replacement? What about without replacement?

7. What is the probability that first king is the third card I draw (with replacement)?

8. In a class of 40 males and 60 females, I give out 4 awards randomly. What is the probability that females will win all 4 awards if the awards must go to different people? What about if the same person can win multiple awards?

9. In a class of 50 males and 90 females, I give out 4 awards randomly. What is the probability that females will win 2 awards if the awards must go to different people? What about if the same person can win multiple awards?

10. At Berkeley, 3/4 of the population is undergraduates. I cold call someone at random and ask for their age. What is the probability that I have to call 10 people before I call an undergraduate? What is the probability that I call 4 undergraduates out of 10 people I call (if I can call someone more than once)?

11. In a dorm of 100 people, there are 20 people who are underage. I go to a party with 40 people. What is the probability that there is at least one underage person there?

12. I roll a fair 6-sided die over and over again until I roll a 6. What is the probability that it takes me more than 10 tries?

13. The number of chocolate chips in a cookie is Poisson distributed with an average of 15 chocolate chips. What is the probability that you pick up a cookie with only 10 chocolate chips in it?

14. The number of errors on a page is Poisson distributed with approximately 1 error per 50 pages of a book. What is the probability that a novel of 300 pages contains at most 1 error?

15. I roll two fair 6 sided die. What is the expected value of their product?

16. While pulling out of a box of cookies, what is the expected number of cookies I have to pull out before I pull out an oatmeal raisin if 25% of cookies are oatmeal raisin and I choose with replacement? What is the variance?

17. What is the expected number of aces I have when I draw 5 cards out of a deck?

18. In a safari, safari-keepers have caught and tagged 300 rhinos. On a safari, out of the 15 different rhinos you see, there are 5 of them expected to be tagged. How many rhinos are there at the safari?

1.4 Hypothesis Testing

19. Chip bags say that they have 14 ounces of chips inside with a standard deviation of 0.5 ounces. You weigh 100 bags and get an average of 13.8 ounces. Can you say that they are underselling you with significance level $\alpha = 0.05$?

20. You flip a coin 100 times and get 60 heads. Can you say the coin isn't fair with significance level $\alpha = 0.05$?

21. You roll a 10 sided die 100 times and get 5 16 times. Can you say the die is biased towards 5 with significance level $\alpha = 0.05$?

22. You take 400 cards and get 107 spades, 112 hearts, 75 diamonds, and 106 clubs. Can you say that the suits are not evenly distributed with $\alpha = 0.05$?

23. In a M&M bag, you get 12 brown ones, 13 yellow ones, 12 red, 14 green, 23 blue, and 22 orange. Is it possible that the colors are evenly distributed with a significance level of $\alpha = 0.05$?

24. You want to know whether living location is independent of a students grade. You interview some people and get the following results:

	Sophomore	Junior	Senior
South	20	30	50
Downtown	10	30	60
North	10	40	50

Are they independent?

25. You are wondering whether performing well in this course and gender are related and you get the following table. Are they related?

	Male	Female
Pass	315	485
Fail	85	115

1.5 Recursion Equations

26. Solve the recurrence relation $a_n = 3a_{n-1} + 2$ with $a_0 = 0$.

27. Verify that $a_n = \binom{2n}{n}$ is a solution to $a_n = \frac{4n-2}{n} a_{n-1}$.

28. Solve the recurrence relation $a_n = 3a_{n-1} + 4a_{n-2}$ with $a_0 = 3$ and $a_1 = 2$.

29. Solve the recurrence relation $a_n = 5a_{n-1} - 4a_{n-2}$ with $a_0 = 3$ and $a_1 = 6$.

30. Find A, B such that $a_n = An + B$ is a solution to the recurrence relation $a_n = 2a_{n-1} + n$.

1.6 Differential Equations

31. Find a solution to $xy' + y = 3x^2$ with $y(1) = 5$.

32. Find the solutions to $y' = y\tan(x) - \sec(x)$.

33. Find the solutions to $e^x y' + y = 1$ with $y(0) = 1$.

2 True/False

34. True False To find $P(B|A)$, it suffices to know just $P(A|B)$ and how to apply Bayes' Theorem.

35. True False Among other things, the proof of Bayes' Theorem for finding $p(B|A)$ depends on being able to split the probability $p(A)$ as a sum probabilities $p(A \cap B)$ and $p(A \cap \overline{B})$, and then further rewrite these as products of certain other probabilities.

36. True False The extra shortcut formula $p(B|A) = \frac{1}{1+\frac{p(A|\overline{B})\cdot p(\overline{B})}{p(A|B)\cdot p(B)}}$ works in one particular case when the standard formula for $p(B|A)$ in Bayes' Theorem fails.

37. True False If a winner in a bicycle race tests positive for steroids, and this test has a very high "True Positive" rate and hence a very low "False positive" rate, then we should take away the winning cup from the athlete because it is extremely likely that he/she has used steroids.

38. True False Error 1 in Hypothesis Testing (reject the null-hypothesis that the person is healthy when the person is actually healthy) is analogous to Testing positive for steroids (event T), yet not having used steroids (event $\overline{S}$); in other words, the significance α corresponds to $p(T \cap \overline{S})$.

39. True False Error 2 in Hypothesis Testing (keep the null-hypothesis that the person is healthy but the person is, in fact, sick) is analogous to Testing negative for steroids (event $\overline{T}$), yet having used steroids (event S); in other words, the power of a test $1 - \beta$ corresponds to $1 - p(\overline{T} \cap S)$.

40. True False To partition a set Ω into a disjoint union of subsets $B_1, B_2, \ldots, B_n$, means that the intersection of these sets is empty; i.e., $B_1 \cap B_2 \cap \cdots \cap B_n = \emptyset$.

41. True False Two disjoint events could be independent, but two independent events can never be disjoint.

42. True False If a fair coin comes up Heads six times in a row, it is more likely that it will come up Tails than Heads on the 7th flip.

43. True False Contrary to how we may use the word "dependent" in everyday life; e.g., event A could be dependent on event B, yet event B may not be dependent on event A; in math "dependent" is a symmetric relation; i.e., A is dependent with B if and only B is dependent with A.

44. True False If A and B are independent events, their complements are also bound to be independent, and to prove this we need a general argument since an example is not sufficient here.

45. True False If A and B are independent events, $\overline{A}$ and B may fail to be independent, but to prove this we need just one counterexample, not a general proof.

46. True False If any pair of events among $A_1, A_2, ..., A_n$ are independent, then all events are independent.

47. True False A random variable (RV) on a probability space (Ω, P) is a function $X : \Omega \to \mathbb{R}$ that satisfies certain rules and is related to the probability function P.

48. True False A RV X could be the only source of data for an outcome space Ω and hence could be very useful in understanding better X 's domain.

49. True False The notation "$X \in E$" means that the RV X starts from event $E \subseteq \Omega$ and lands in $\mathbb{R}$.

50. True False The notation "$X^{-1}(E)$" for a RV X and event $E \subseteq \Omega$ means to take set $B \subseteq \Omega$ of the reciprocals of all elements in E that are in the range of X .

51. True False The PMF of a RV X on probability space (Ω, P) is a third function $f : \mathbb{R} \to [0,1]$ such that the composition of X followed by f on any $\omega \in \Omega$ is equal to P; i.e., such that $f(X(\omega)) = P(\omega)$.

52. True False It is possible that $f(x) > P(x)$ for some $\omega \in \Omega$ and the corresponding $x = X(\omega) \in \mathbb{R}$ where X a discrete RV on (Ω, P) with PMF f .

53. True False To show that two RV's $X, Y : \Omega \to \mathbb{R}$ are independent on (Ω, P) , we can find two subsets $E, F \subseteq \mathbb{R}$ for which $P(X \in E \text{ and } Y \in F) = P(X \in E) \cdot P(Y \in F)$.

54. True False Two Bernoulli trials are independent only if the probability of success and failure are each $\frac{1}{2}$.

55. True False The product X and the sum Y of the values of two flips of a fair coin (H=1, T=0) are dependent random variables.

56. True False To turn the experiment of "rolling a die once" into a Bernoulli trial, we need to split its outcome space into two disjoint subsets and declare one of them a success.

57. True False Several Bernoulli trials performed on one element at a time from a large outcome space Ω, without replacement, are approximately independent because what happens in one Bernoulli trial hardly affects the ratio of "successes" to "failures" in the remainder of the population.

58. True False The probability of having 20 women within randomly selected 40 people is about 50%, assuming that there is an equal number of women and men on Earth.

59. True False The hypergeometric distribution describes the probability of k "successes" in n random draws without replacement from a population of size N that contains exactly m "successful" objects.

60. True False While the hypergeometric and binomial probabilities depend each on 3 parameters and 1 (input) variable, the Poisson probability depends only on 1 parameter and 1 (input) variable.

61. True False To approximate well the probability of k successes in a large number n of independent Bernoulli trials, each with individual probability of success p that is relatively small, we can use the formula $\frac{(np)^k e^{-np}}{n!}$.

62. True False $E(X-Y) = E(X)-E(Y)$ for any R.V.s X and Y, regardless of whether they are independent or not.

63. True False $Var(X) = E(X^2) - E^2(X)$ holds true because, essentially, the expected value has linearity properties.

64. True False Splitting a R.V. X as a sum of simpler R.V.'s X_i could be advantageous when we want to compute $Var(X)$, but we need to be careful that these X_i's are independent.

65. True False If X is the geometric R.V., then $Var(X) = \frac{1-p}{p^2}$.

66. True False If X is the hypergeometric R.V. in variable k and with parameters m, n, N, then $N = \frac{mn}{E(X)}$.

67. True False If Santa Claus randomly throws n presents into n chimneys (one present per chimney), on the average one home will receive their intended present, regardless of how large or small n is.

68. True False For any independent R.V's X and Y, we have $Var(aX + bY) = a^2Var(X) + b^2Var(Y)$, but $Var(XY) = Var(X) \cdot Var(Y)$ only if $E(X) = 0 = E(Y)$ or $X = 0$ or $Y = 0$ or both X and Y are constants.

69. True False As long as several RV's $X_1, X_2, ..., X_n$ are identically distributed, their average RV $\overline{X}$ will have the same mean as each of them, but the standard error of $\overline{X}$ may or may not be equal to $\frac{SE(X_3)}{\sqrt{n}}$.

70. True False The formula for the mean of the average $\overline{X} = \frac{X_1+X_2+\cdots+X_n}{n}$ of independent identically distributed RV's has $\sqrt{n}$ in the denominator.

71. True False The formula for the variance $Var(X + Y) = Var(X) + Var(Y)$ works regardless of whether the RV's X and Y are independent or not.

72. True False To approximate the height of a tall tree (using similar triangles and measurements along the ground – without climbing the tree!) it is better to ask several people to do it independently of each other and then to average their results, than to do it once just by yourself.

73. True False According to the Central Limit Theorem, the normalized distribution $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ is the standard normal distribution for n large, where $\overline{X}$ is the average of n independent, identically distributed variables, each with mean μ and standard error $\sigma/\sqrt{n}$.

74. True False If we have data $\{x_1, x_2, \ldots, x_N\}$ from the whole population, we should divide by n instead of $(n-1)$ in the formula for the sample standard deviation.

75. True False The smaller the 95% confidence interval is, the lower our confidence is that the true parameter is in that interval.

76. True False To reject the null hypothesis H_0 when it is actually true is equivalent to making a Type 2 error with significance α.

77. True False The null hypothesis is a theory that we believe is true.

78. True False The higher the significance of a test, the higher the probability of rejecting a true null hypothesis.

79. True False Adding up the power and the significance of a test yields 1.

80. True False A type-2 error made by a road patrol may result in letting drunken drivers continue driving.

81. True False The significance of a test shows how often, on the average, we can make a Type 1 error.

82. True False Using two-sided Alternative Hypotheses H_1 may lead to twice as large significance as their one-sided analogs.

83. True False The p-value of a possible test result r is the probability that the experiment produces a result that is equally or more extreme (towards H_1) than r, assuming H_0 is true.

84. True False In class we concluded that marital status and employment status for men in the age group 35-44 years old must be dependent since the resulting r-value of the test statistic $R = \sum_{i,j} \frac{(N_{ij}-M_{ij})^2}{M_{ij}}$ came to $r \approx 31.61 > \alpha = 0.05$.