# Sample Exam Questions 2 Solutions – Math 263 (sect 9)

1. If you draw an M&M candy at random from a bag of the candies, the candy you draw will have one of six colors. The probability of drawing each color depends on the proportion of each color among all candies made. Assume the table below gives the probabilities for the color of a randomly chosen M&M:

| Color | Brown | Red | Yellow | Green | Orange | Blue |
|---|---|---|---|---|---|---|
| Probability | 0.3 | 0.3 | ? | 0.1 | 0.1 | 0.1 |

a. What is the probability of drawing a yellow candy? **0.1**

b. What is the probability of not drawing a red candy? **0.7**

c. What is the probability that you draw neither a brown nor a green candy? **0.6**

d. If you select two M&M's and the colors are independent, then what is the probability that the two M&M's are the same color?

**0.3*0.3 + 0.3*0.3 + 0.1*0.1 + 0.1*0.1 + 0.1*0.1 + 0.1*0.1= 0.22**

2.. Ignoring twins and other multiple births, assume babies born at a hospital are independent events with the probability that a baby is a boy and the probability that a baby is a girl both equal to 0.5.

a.. What is the probability that the next three babies are of the same sex? **0.25**

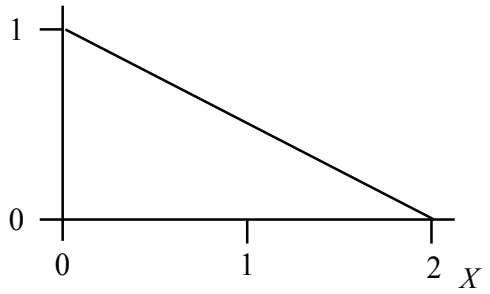
b. Define events $A$ = {the next two babies are boys} and $B$ = {at least one of the next two babies is a boy}. What do we know about events $A$ and $B$?
**(D) none of the above.**

c. Define event $B$ = {at least one of the next two babies is a boy}. What is the probability of the complement of event $B$? **0.25**

d. What is the probability that at least one of the next three babies is a boy?
**1-0.125=0.875**

3. For each of the following random variables is it discrete or continuous?
a. number of text messages a random student sends in a month **discrete**
b. a random variable with a normal distribution **continuous**
c. the weight of a randomly chosen student **continuous**
d. the volume of gasoline lost due to evaporation during the filing of a gas tank **continuous**
e. number of free throws made by a basketball player in 10 attempts. **discrete**

4.. Suppose $X$ is a continuous random variable taking values between 0 and 2 and having the probability density function below:




a. What is $P(1 \leq X \leq 2)$? **0.25**

b. What is P(X>1.5|X>1) ? **=P(X>1.5)/P(X>1)=0.25**

5..The American Veterinary Association claims that the annual cost of medical care for dogs averages \$100 with a standard deviation of \$30. The cost for cats averages \$120 with a standard deviation of \$35. Let X be the cost of a randomly chosen dog's medical care minus the cost of a randomly chosen cat's medical care.

a. Find the mean and standard deviation of X . **mean= -20, variance is sum of variance for dog and variance for cat. So standard dev=46.1**

b. If the difference in costs follows a Normal distribution, what is the probability that the cost for someone's dog is higher than for the cat? **This is the same as asking for the probability that X is positive. Using calculator or tables this is 0.3336.**

6.. The following table describes the probability distribution for the random variable X that counts the number of times a customer visits a grocery store in a 1-week period:

| Visits | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P$(Visits) | 0.1 | 0.25 | 0.3 | **p** | 0.1 |

a. What is the value of **p** ? **p=0.25**

b. Find the mean of X. **mean=2**

c. Find the standard deviation of X. **var=1.3, std dev=1.14**

7. Among U of A freshmen, 40% are taking one or more math courses, 30% are taking one or more English courses and 40% are taking one or more arts courses. 8% are taking one or more math courses and one or more arts courses. 8% are taking one or more English courses and one or more arts courses. There are no students taking courses in all three of the areas of math, English and arts. 12 % are taking one or more English courses and one or more math courses.
a. If we pick a freshman at random what is the probability he or she is not taking any course in the areas of math, English or arts? **0.18**
b. Are the events taking one or more math courses and taking one or more English courses

independent? Explain your answer. **Probability of their intersection is 0.12, product of their individual probabilities is 0.4*0.3=0.12. These are equal so they are independent.**
c. Are the events taking one or more arts courses and taking one or more English courses independent? Explain your answer. **Probability of their intersection is 0.08, product of their individual probabilities is 0.4*0.3=0.12. These are not equal so they are not independent.**
d. We pick a student at random and ask whether he or she is taking one or more math courses but not whether he or she is taking any English courses. If the student is taking one or more math courses what is the probability the student is taking one or more English courses?
**We want P(one or more English| one or more math). The intersection has probability 0.12 and the given event has probability 0.40. So answer is 0.12/0.40=0.3.**

8. The weight of medium-size tomatoes selected at random from a bin at the local supermarket is a random variable with mean $\mu = 10$ oz and standard deviation $\sigma = 1$ oz. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable Y to be the weight of the bag containing the four tomatoes.

a. What is the mean of Y? **4 * 10 oz = 40 oz**

b. What is the standard deviation of Y? **variance = 4* 1 = 4, so stand dev = 2 oz**

9. 22% of the population of Arizona is over 60 years old. Among the people over 60, 45 % are retired. Among the people 60 and younger only 8% are retired.

a. If I pick an Arizona resident at random what is the probability he or she is retired?
**0.22*0.45 + 0.78*0.08=0.1614**
b. I pick an Arizona resident at random and find out he or she is retired. What is the probability he or she is over 60?
**P(over 60 | retired) = P(over 60 and retired)/P(retired)=0.22*0.45/0.1614 = 0.6134**

10. When figure skaters need to find a partner for "pair figure skating," it is important to find a partner who is compatible in weight. The weight of figure skaters can be modeled by a Normal distribution. For male skaters, the mean is 170 lbs with a standard deviation of 10 lbs. For female skaters, the mean is 110 lbs with a standard deviation of 5 lbs. Let the random variable $X$ = the weight of female skaters and the random variable $Y$ = the weight of male skaters.

a. The weight of a pair of figure skaters (a male and a female) can be thought of as a new random variable. Let the random variable $W = X + Y$. What is the mean of this new random variable $W$?
**Mean = 170+110=280 lbs**

b. Suppose we consider the weights of the male partner and the female partner to be independent. What is the standard deviation of the random variable $W$?

**Variance=10*10+5*5=125, so standard deviation=sqrt(125)=11.2 lbs**

c. It does seem likely that the weights of the male and the female partner would be independent. If the correlation $\rho$, between $X$ and $Y$ equals 0.77, what is the standard deviation of the random variable $W$

**Variance = 10*10+5*5 + 2 * 0.77 * 10 * 5=202, so standard deviation = sqrt(202)=14.2 lbs**

11. In humans the blood platelet count (measured in units of 1,000) has a mean of 268 with a standard deviation of 166. We choose a random sample of 110 people and compute the sample mean.

a. What is the mean and standard deviation of this sample mean? **Mean is 268, standard deviation is 166/sqrt(110)=15.83.**

b. What is the probability the sample mean is within 25 of the population mean? **To be within 25 of the poplution mean the sample mean must be between 243 and 293. So we need the probability a normal RV with the mean and standard deviation from part a lies between 243 and 293. Normcdf() gives 0.8857.**

c. Suppose I want to choose a large enough sample that the standard deviation of the sample mean is only 10. How big a sample do I need to use? **For a sample of size n the standard deviation of the sample mean is 166/sqrt(n). This must be 10. So sqrt(n)=166/10=16.6. So n is 16.6*16.6=276 (I rounded up since n must be an integer.)**

d. Now suppose we only sample 5 people and use the normcdf() function in our calculator to estimate the probability that the sample mean is greater than 280. Is this justified and if not what additional information would I need to justify it**? We need to know that the distribution of the sample mean is approximately normal. If the sample were large this would follow from the central limit theorem. But the sample only has size 5, so the central limit theorem does not apply. So this is justified only if we know that the population distribution is normal.**

12. In order to determine if smoking causes cancer, researchers surveyed a large sample of adults. For each adult they recorded whether the person had smoked regularly at any period in their life and whether the person had cancer. They then compared the proportion of cancer cases in those who had smoked regularly at some time in their lives with the proportion of cases in those who had never smoked regularly at any point in their lives. The researchers found a higher proportion of cancer cases among those who had smoked regularly than among those who had never smoked regularly. What type of study is this?

**A) An observational study.**

13. Many studies are trying to find a cure for chronic back pain. In one such study, a physician is comparing the medication currently being used (drug A) to a newly developed drug (drug B). Seventy-three volunteers, suffering from chronic back pain, are participating in this study. The physician's assistant has a list of all 73 subjects and randomly divides the subjects into two groups. Group 1 will receive drug A and Group 2 will receive drug B. The assistant is the only one who knows to which group the subjects have been assigned. The physician monitors the subjects over a 2-month period and the amount of improvement is recorded. What type of study is this?

**C) A double-blind experiment.**

14. In order to assess the effects of exercise on reducing cholesterol, a researcher sampled 50 people from a local gym who exercise regularly and 50 people from the surrounding community who, were assumed, do not exercise regularly. Each subject reported to a clinic to have their cholesterol measured. The subjects were unaware of the purpose of the study, and the technician measuring the cholesterol was not aware of whether the subject exercises regularly or not. What type of study is this?

**A) An observational study.**

15. A study is designed to determine whether grades in a statistics course could be improved by offering special review material. The 250 students enrolled in a large introductory statistics class are also enrolled in one of 20 lab sections. The 20 lab sections are randomly divided into 2 groups of 10 lab sections each. The students in the first set of 10 lab sections are given extra review material during the last 15 minutes of each weekly lab session. The students in the remaining 10 lab sections receive the regular lesson material, without the extra review material. The grades of the students who reviewed weekly were higher, on average, than the students who did not review every week. What type of study is this?

**B) An experiment, but not a double-blind experiment.**

16. A market research company wishes to find out which of two Internet search engines the population of students at a university prefers to use: Google or MSN Search. A random sample of students is selected, and each one is asked to search for a certain subject using Google and then MSN, or vice versa. The order of the two searches was determined at random. They then indicate which Internet search engine they prefer. What type of study is this?

**B) An experiment, but not a double-blind experiment.**

17. One hundred volunteers who suffer from severe depression are available for a study. Fifty are selected at random and are given a new drug that is thought to be particularly effective in treating severe depression. The other 50 are given an existing drug for treating severe depression. A psychiatrist evaluates the symptoms of all volunteers after four weeks in order to determine if there has been substantial improvement in the severity of the depression.

a. What is the explanatory variable or factor in this study?

**A) Which drug the volunteers receive.**

b. In which situation would this study be double-blind?

**D) All of the above.**

18. Suppose volunteers were first divided by gender, and then half of the men were randomly assigned to the new drug and half of the women were assigned to the new drug. The remaining volunteers received the other drug. What is this an example of?

**C) A block design.**

19. Olivia is planning to take a foreign language class. To research how satisfied other students are with their foreign language classes, she decides to take a sample of 20 such students. The university offers classes in four languages: Spanish, German, French, and Japanese. She will select a simple random sample of five students from each language. What sampling technique is using ?

**B) A stratified sample.**