MST 109 - Midterm 2 - Fall 2017

Elementary Probability and Statistics

November 2, 2017

Instruction:

- You have 50 minutes to finish this exam.
- During this exam you are allowed to look at your notes and use your computers.
- It is expected that each student during this exam will conduct himself, herself or themselves within the guidelines of the WFU Honor Code. All academic work should be done with the high level of honesty and integrity that the university demands.

Name: Solution

Section: Section G (1:00 pm- 1:50 pm) Section H (2:00pm - 2:50 pm)

I) An auto insurance company has 10,000 policyholders. Each policyholder is classified as

i) young or old;

ii) male or female;

Of these policyholders, 4000 are young, 6400 are male. The policyholders can also be classified as 1800 young males.

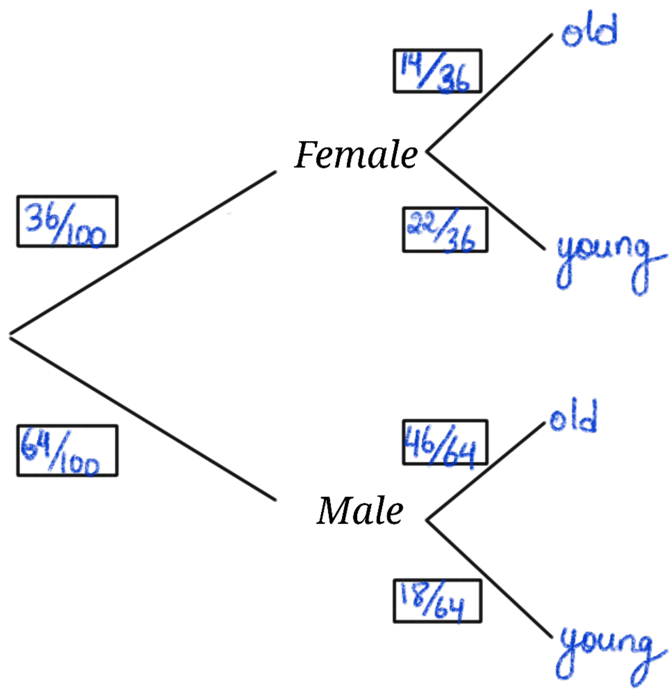
1. Find the probability that a randomly chosen policyholder is old.

$$P(\text{old}) = \frac{6000}{10000} = 6/10 = 3/5$$

2. Knowing that a policyholder is a young person, what is the probability that they are male?

$$P(\text{male}/\text{young}) = \frac{1800}{4000} = \frac{18}{40} = \frac{9}{20}$$

3. Complete the following tree diagram with the correct probabilities.



4. Knowing that the policyholder is a female, what is the probability that she is old?

$$P(\text{old}/\text{Female}) = 14/36$$

II)A doctor is studying the relationship between blood pressure and heartbeat abnormalities in her patients. She tests a random sample of her patients and notes their blood pressures (high, low, or normal) and their heartbeats (regular or irregular). She finds that:

(i) 14% have high blood pressure. (HBP)

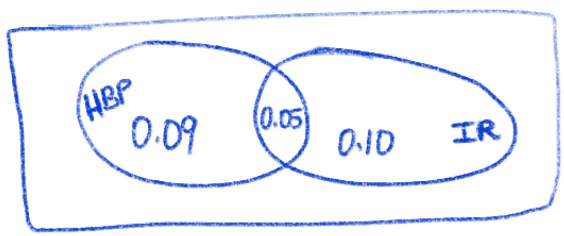
(ii) 15% have a irregular heartbeat. (IR) [regular heartbeat : (R)]

(iii) Of those with an irregular heartbeat, one-third have high blood pressure.

1. What is the percentage of patients that have high blood pressure and an irregular heartbeat?

$$P(HBP \text{ and } IR) = P(HBP/IR) \times P(IR)$$
$$= \frac{1}{3} \times 0.15 = 0.05$$

2. Draw a Venn-diagram showing the percentage of patients with high blood pressure, patients with irregular heartbeat and the ones with both.



If in addition we know that:

(iv) 22% have low blood pressure. (LBP)

(v) Of those with normal blood pressure (NBP), one-eighth have an irregular heartbeat.

3. What is the proportion of the patients that have normal blood pressure?

$$P(NBP) = 0.64$$
$$= 1 - P(HBP) - P(LBP)$$

4. Calculate the portion of the patients selected who have a regular heartbeat and low blood pressure.

$$P(NBP \text{ and } IR) = P(IR/NBP) \times P(NBP)$$
$$= \frac{1}{8} \times 0.64$$
$$= 0.08$$
$$\Rightarrow P(IR \text{ and } LBP) = 0.15 - 0.08 - 0.05$$
$$= 0.02$$
$$\Rightarrow P(R \text{ and } LBP) = 0.20$$

III) A lab network consisting of 200 computers was attacked by a computer virus. This virus enters each computer with probability 0.4, independently of other computers.

1. Find the probability that the virus enters exactly 10 computers.

   Binomial dist $n = 200$ $p = 0.4$

   $P(X = 10) = \text{Binom.dist}(10, 200, 0.4, \text{false})$

   $= 0$

2. Find the probability that the virus enters at least 50 computer.

   $P(X \geq 50) = 1 - P(X \leq 49)$

   $= 1 - \text{Binom.dist}(49, 200, 0.4, \text{true})$

   $= 0.999997139 \sim 1$

3. **Check** that you can approximate this Binomial probability with a Normal distribution and calculate the mean and standard deviation.

   $np = 80 \geq 10$ $n(1-p) = 120 \geq 10 \Rightarrow$ Yes we can approximate

   mean $= 80$ s.d. $= \sqrt{np(1-p)} = 6.9282$

4. Using the Normal approximation calculate the probability that the virus entered fewer than 95 computers.

   $P(X < 95) = P(X \leq 94)$

   $= \text{Norm.dist}(94, 80, 6.9282, \text{true})$

   $= 0.9783$

IV) The distribution of results from a cholesterol test has a mean of 180 and a standard deviation of 20. A sample size of 40 is drawn randomly.

Find the probability that the sum of the 40 values is greater than 7,600.

$$s.d. = \frac{20}{\sqrt{40}} = 3.1623$$

$$P(\text{Sum} \leq 7600) = P\left(\bar{x} \leq \frac{7600}{40}\right)$$

$$\Rightarrow P(\text{Sum} \geq 7600) = 1 - P(\text{Sum} < 7600)$$

$$= 1 - \text{Norm.dist}\left(\frac{7600}{40}, 180, 3.1623, \text{true}\right) = 0.000782761$$

V) Sulfur compounds cause "off-odors" in wine, so winemakers want to know the odor threshold, the lowest concentration of a compound that the human nose can detect. The odor threshold for dimethyl sulfide (DMS) in trained wine tasters is about 25 micrograms per liter of wine ($\mu g/l$). The untrained noses of consumers may be less sensitive, however. Here are the DMS odor thresholds for 10 untrained students.

31	31	41	37	22	34	32	31	20	23

1. Assume that the standard deviation of the odor threshold for untrained noses is known to be $\sigma = 8\mu g/l$. State two other conditions that need to be satisfied so we can carry on a confidence interval analysis.

The $\underline{\underline{3}}$ remaining conditions are

- You need an SRS
- You need the population to have a normal dist.
- You need the mean to be unknown.

2. Find the 95% confidence interval.

then the 95% conf. interval is given by

$$\left[\bar{x} - z^* \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \frac{\sigma}{\sqrt{n}}\right] = [25.2415, 35.1585]$$

$z^* = 1.96$ $\bar{x} = 30.2$

So $\bar{x} - z^* \frac{\sigma}{\sqrt{n}} = 25.2415$

$\bar{x} + z^* \frac{\sigma}{\sqrt{n}} = 35.1585$

VI) Suppose the mean recovery time for individuals who do not use any form of treatment is generally thought to be 30 days with standard deviation equal to 8. A pharmaceutical company manufacturing a certain cream wishes to determine whether the cream shortens, extends (i.e. Changes), or has no effect on the recovery time. The company chooses a random sample of 100 individuals who have used the cream, and determines that the mean recovery time for these individuals was 28.5 days.

1) Write down your null hypothesis and alternative hypothesis.

$H_0: \mu = 30$

$H_a: \mu \neq 30$

2) Carry on a significance test, and state weather the sample is significant at at 5% significance level.

$$Z_{statistic} = \frac{28.5 - 30}{\sigma/\sqrt{n}} = \frac{1.5}{8/10} = -1.875$$

$$P_{value} = 2\,P(z > |-1.875|) = 2P(z > 1.875)$$

$$= 2\left[1 - \text{Norm.dist}(1.875, 0, 1, \text{true})\right] = 0.0608 \geqslant 0.05$$

This Sample is not significant, So we don't have enough evidence to reject H_0.

VII) Choose the best answer:

1) Suppose the P-value for a hypothesis test is 0.270. Using $\alpha = 0.05$, what is the appropriate conclusion?

(a) Reject the null hypothesis.

(b) Reject the alternative hypothesis.

(c) Do not reject the null hypothesis.

(d) Do not reject the alternative hypothesis.

2) The body mass index (BMI) of all American young women follows a Normal distribution with standard deviation $\sigma = \mathbf{7.5}$. How large a sample would be needed to estimate the mean BMI μ in this population with a margin of error equals **1** and with **95%** confidence? (Use $z^* = 1.96$ for 95% confidence.)

(a) 152

(b) 216

(c) 245

(d) cannot be calculated from the information given