# STAT 614 -HW2

# Problem 1

Vegetarians have fewer heart attacks than non-vegetarians of the same age and sex. One reason is that vegetarians have (on average) lower blood pressure. But does a vegetarian diet actually cause lower blood pressure? The following studies were used to explain the difference in blood pressure between vegetarians and non-vegetarians.

**Study 1**: Surveys about diet and health were sent to a random sample of 500 adults in a northwestern city. Of the 233 respondents, those who chose to be vegetarians tended to have lower blood pressure than non-vegetarians.

**Study 2**: A sample of 200 adults who are not vegetarians were recruited by newspaper advertisement to participate in a study. Blood pressure was measured in all subjects at the start of the study. Half were randomly assigned to eat a vegetarian diet and the other half were assigned to their regular diet. After a year, blood pressure was measured again in each subject and the vegetarian group had, on average, a larger decrease in their blood pressure.

**Study 3**: Trappist monks are strict vegetarians, while Benedictine monks follow a more standard Western diet. A study of these two groups revealed that Trappist monks have (on average) lower blood pressure than Benedictine monks.

Answer the following questions about the above studies:

1. Which of these studies, if any, is an example of a randomized experiment?

**Ans**: Study 1&2

1. In which of these studies, if any, were subjects randomly chosen to participate?

**Ans:** Study 1

1. Which study provides the strongest evidence that a vegetarian diet causes lower blood pressure? Explain your reasoning.

**Ans**: I think Study 2 provides the strongest evidence that a vegetarian diet causes lower blood pressure. In the test, a randomly selected group are divided into two and assigned to eat on designated diets, which minimize the bias in both groups and further been tested that vegetarian group had decrease in their blood pressure.

1. For the two remaining studies (that you did not select for part c), pick one and briefly describe why you did not select it in part c.

**Ans**: In study 1, there are still lots of possible reasons will affect the blood pressure results, such as different vegetarian diet type or their current health concern may also cause low blood pressure.

# Problem 2

From *The Statistical Sleuth, Third Edition*, Chapter 1, problems 19 & 20. Please include a description of the method you used for problem 20.

Note: If you don’t have a copy of the textbook, go to the Information area in Blackboard where the first two chapters are posted (as a very rough pdf).

**Ans**:

1. **problems 19**

* Write down the names and ages

|  |  |
| --- | --- |
| **Name** | **Age** |
| Amy | 22 |
| Betty | 30 |
| Cindy | 57 |
| Doris | 62 |
| Eve | 34 |
| Fiona | 54 |
| Gary | 81 |
| Helen | 21 |
| Iris | 19 |
| Jenny | 17 |

* Using coin flips comes out as

|  |  |
| --- | --- |
| **Group A(Head)** | **Group B(Tail)** |
| Betty | Fiona |
| Helen | Amy |
| Doris | Cindy |
| Iris | Gary |
| Eve |
| Jenny |

* Group B tends to get many of the older subjects.
* There’s no way we can predict which group will have higher average age since the possibility of assigning to group A or B is 50% of random selection

1. **problems 20**

* randomize select, ensure each group ends up five people

|  |  |
| --- | --- |
| **Group A** | **Group B** |
| Jenny | Doris |
| Amy | Betty |
| Eve | Fiona |
| Iris | Cindy |
| Helen | Gary |

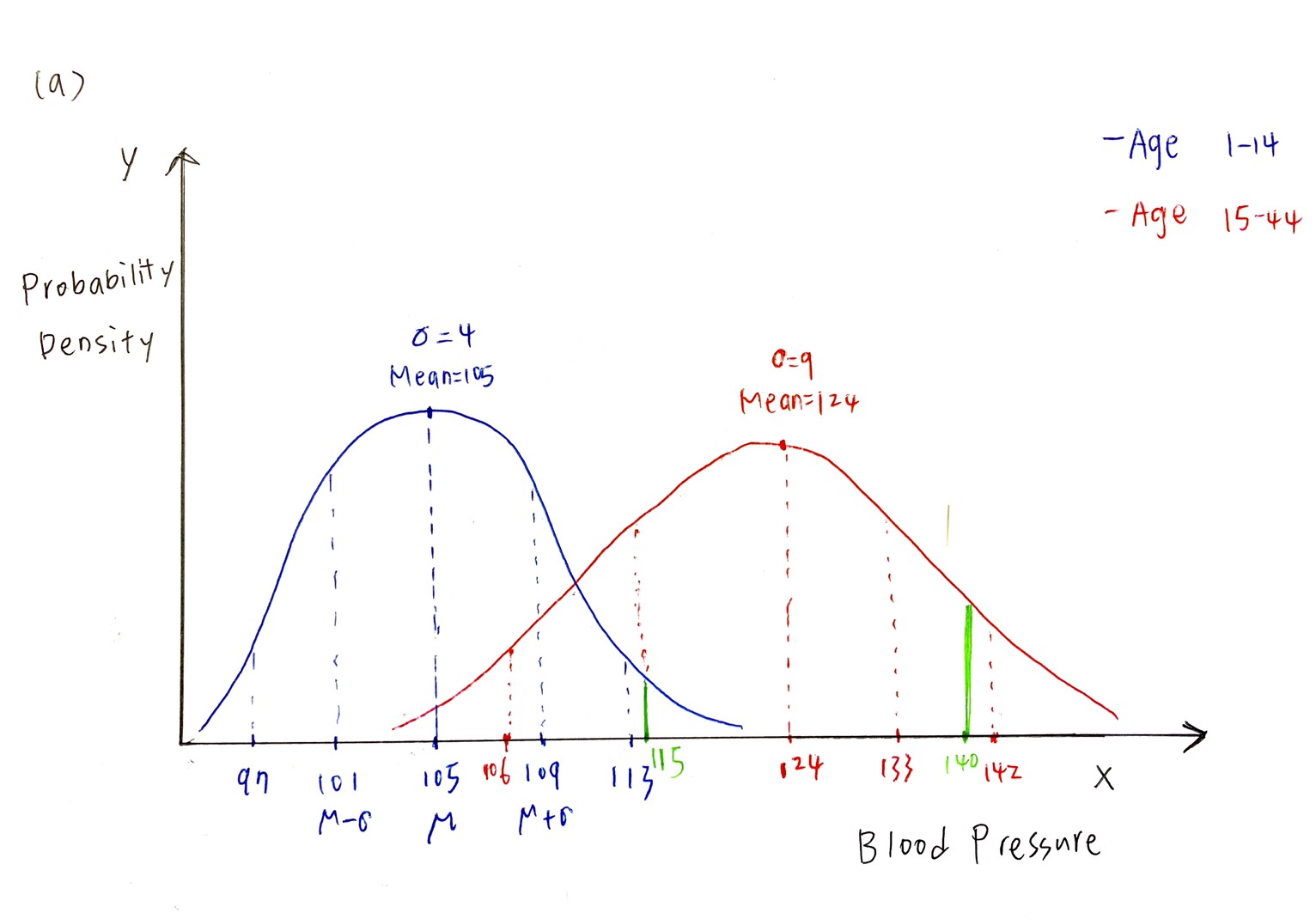
# Problem 3

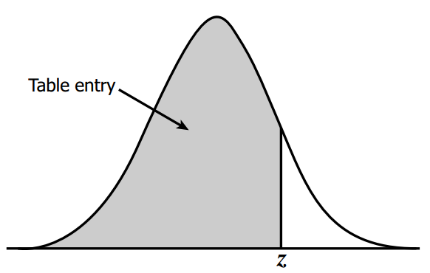
People are classified as hypertensive if their systolic blood pressure is higher than a specified level for their age group, according to the last column of the following table. Assume systolic blood pressure is *Normally distributed* with mean and standard deviation given in the table for the age groups 1-14 and 15-44, respectively.

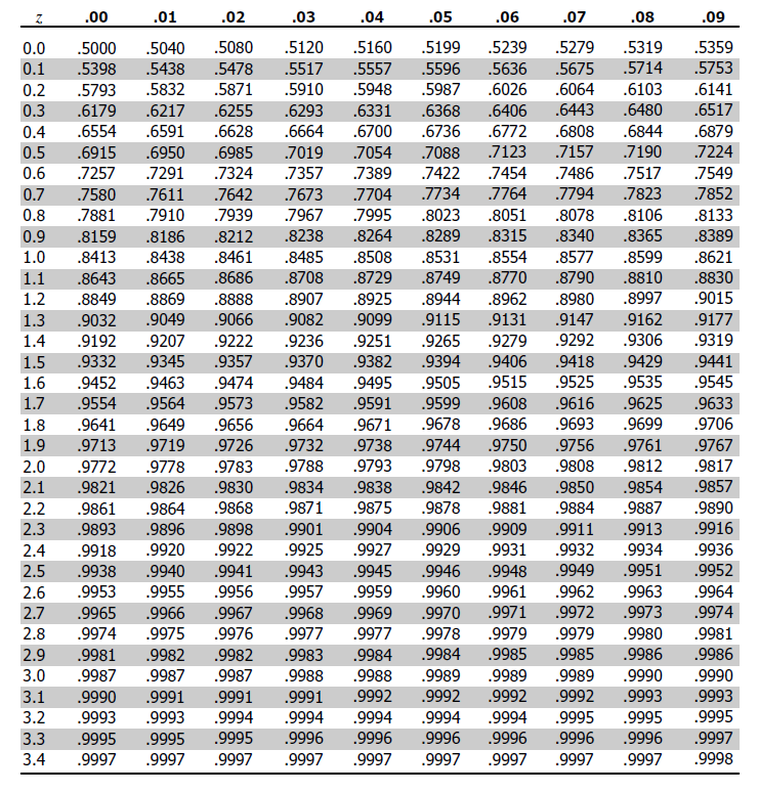
|  |  |  |  |
| --- | --- | --- | --- |
| Age group | Mean | St. Dev | Level |
| 1-14 | 105.0 | 4.0 | 115.0 |
| 15-44 | 124.0 | 9.0 | 140.0 |

1. Sketch a graph of the distribution of systolic blood pressure for both groups (on the same plot). Label the x- and y-axes, the mean, the standard deviation, and indicate the hypertensive cut-off for each group. You may take a picture of your graph and copy-and-paste it into your homework to include it with your solution. DO NOT submit a separate image file! Paste it into your homework document (and I recommend saving your homework document as a pdf to make sure it is readable in Blackboard).

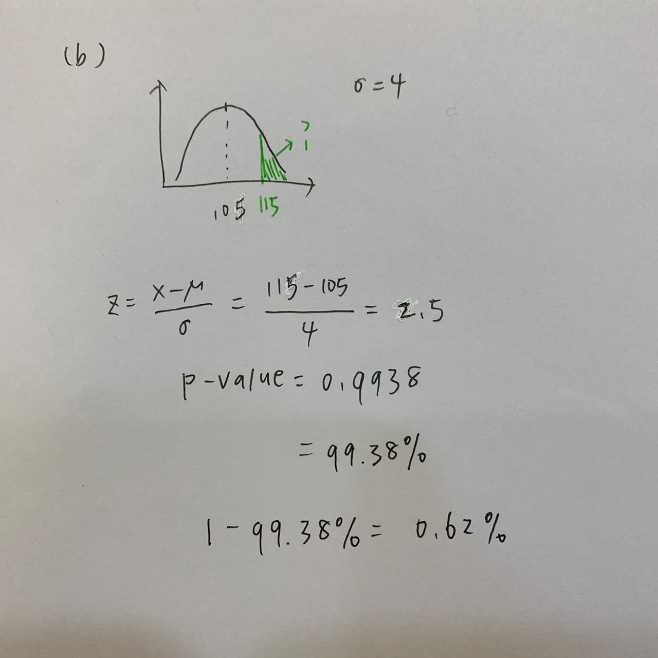
Ans:



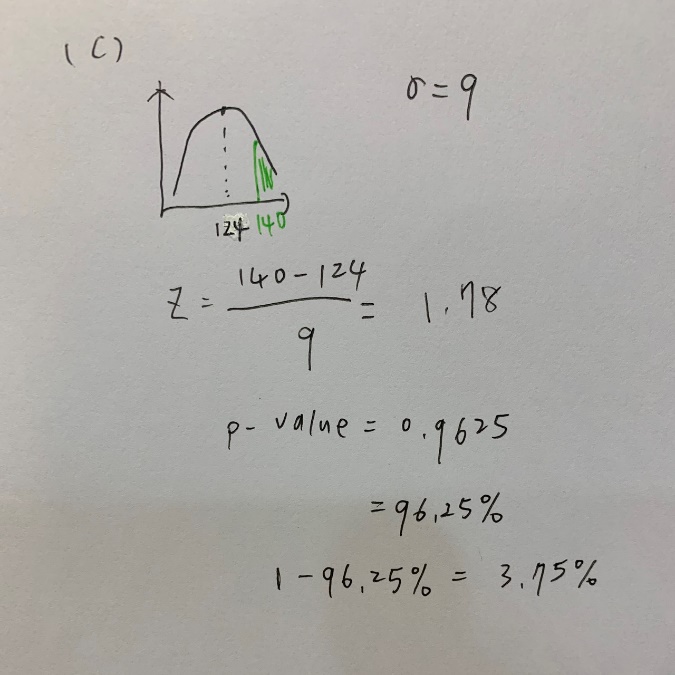
 Find values on the right of the mean in this z-table.



1. What proportion of 1- to 14-year-olds are hypertensive?

**Ans: 0.62%**  / 1-pnorm(115,105,4)= 0.62%

1. What proportion of 15- to 44-year-olds are hypertensive?

**Ans: 3.75%** / 1- pnorm(140,124,9)= 3.77%

1. What is the probability that a randomly select 1-14-year-old is not hypertensive?

**Ans**: **99.38%**

1-0.62% = 99.38%/ pnorm(115,105,4) = 99.38%

1. Suppose two people are selected at random from 1-14-year-old age group. What is the probability that at least one of them is not hypertensive?

**Ans**: 1-P(both hypertensive) = 1-0.62%\*0.62% = **99%**

1. Suppose ten people are selected at random from 1-14-year-old age group. What is the probability that all of them is not hypertensive?

**Ans**: 99.38%^10 = **93.97%**

# Problem 4

Find **one** scholarly article which uses **one** of a normal, Poisson, binomial, or any other distribution of interest to you.

Give the following information from your article:

1. The full citation of the article (authors, title of article, journal name, year, volume, page #s).
2. A brief summary of the purpose of the article based on the information in the abstract. You can quote the article but then also put the purpose in your own words.
3. A brief summary of how the distribution is used in the article. You do not need to go into a lot of detail here. Just try to get a sense of what they are using the distribution for.
4. How you found the article (for example, through a PubMed or Google Scholar search).
5. Why you picked that article.

Note: The next page has an example answer to this problem to give you a sense of what I am looking for with this. (Which also means you cannot use this article for your HW solution!)

**Ans**:

1. Rahman, S. A., West, S. K., Mkocha, H., Munoz, B., Porco, T. C., Keenan, J. D., & Lietman, T. M. (2015). The distribution of ocular Chlamydia prevalence across Tanzanian communities where trachoma is declining. *PLoS neglected tropical diseases*, *9*(3), e0003682.

<https://pubmed.ncbi.nlm.nih.gov/25815466/>

1. “The World Health Organization has targeted trachoma for elimination as a public health concern by the year 2020. Local programs may benefit if a single survey could reveal whether infection was headed towards elimination.” quoted from the article. In this case study, they want to find out whether the program of trachoma elimination is working on track or not.
2. The authors indicate that “Models correctly predict that infection prevalence across communities where a disease is disappearing is best described by an exponential distribution.”, they further state that an exponential distribution gave the best fit to prevalence data.
3. PubMed, “exponential distribution” and “World Health Organization”
4. Because of the covid-19 pandemic, I am kind of interested in what other disease is also be included in public health concern