

**MAT 2379 3X (Spring 2010)**  
**Assignment 2 - Solutions**  
**Total = 20 points**

1. [5 points] A wildlife biologist examines frogs for a genetic trait he suspects may be linked to sensitivity to industrial toxins in the environment. Previous research had established that this trait is usually found in approximately 10% of frogs. He collects and examines twelve frogs. Assuming that a frog will have this trait with a probability of 0.10, answer the following questions.
- (a) Give the expected number of frogs with this trait and the standard deviation for the number of frogs with this trait.
  - (b) Compute the probability that none of the frogs will have the trait.
  - (c) Compute the probability that at least 2 frogs will have the trait.
  - (d) Compute the probability that 3 or 4 frogs will have the trait.
  - (e) Now let us assume that the number of selected frogs is  $n = 300$ . Compute the probability that at most 30 will have the trait.

**Solution:** Let  $X$  be the number of frogs with this trait. The random variable  $X$  follows a binomial distribution with  $n = 12$  and  $p = 0.10$ .

**Note:** We can either use R to compute the probabilities with the binomial distribution or use the p.m.f.

(a)  $E[X] = np = (12)(0.1) = 1.2$  and  $\sigma_X = \sqrt{12(.1)(.9)} = 1.03923$

(b)  $P(X = 0) = {}_{12}C_0(.1)^0(.9)^{12} = 0.28243$

(c)

$$P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)] = 0.3410$$

(d)

$$P(\{X = 3\} \cup \{X = 4\}) = P(X = 3) + P(X = 4) = 0.1065$$

(e)  $P(X \leq 30) = 0.5484$

2. [5 points] Suppose that the age that a child starts to walk can be modeled as a normal random variable with a mean of 13 months and a standard deviation of 1.5 months.

- (a) What is the probability that a child would walk before the age of 11 months?
- (b) What is the probability that a child would walk before the age of 15 months?
- (c) What is the probability that the age that a child would walk is between 11 and 15 months?
- (d) What is the probability that a child would be exactly 13 months when he/she starts walking?
- (e) What is our risk of being wrong if we were to bet that the child would start walking between the ages of 12 and 15 months.
- (f) Find an age (in months) such that 95% of the children would start walking before this age.

**Solution:** Let  $X$  be the age at which the child starts to walk. So  $X \sim N(\mu = 13, \sigma^2 = 1.5^2)$ .

(a)

$$P(X < 11) = \Phi\left(\frac{11 - 13}{1.5}\right) = \Phi(-1.33) = 0.0918$$

Alternatively, we could use R to obtain  $P(X < 11) = 0.09121122$ .

(b)

$$P(X < 15) = \Phi\left(\frac{15 - 13}{1.5}\right) = \Phi(1.33) = .9082$$

Alternatively, we could use R to obtain  $P(X < 15) = 0.9087888$ .

(c)

$$P(11 < X < 15) = P(X < 15) - P(X < 11) = .9082 - 0.0918 = 0.8164.$$

Alternatively, we could use R to obtain  $P(11 < X < 15) = 0.8175776$

(d)  $P(X = 13) = 0$

(e)

$$\begin{aligned}P(X < 12 \text{ or } X > 15) &= 1 - P(12 \leq X \leq 15) \\&= 1 - \left[ \Phi\left(\frac{15 - 13}{1.5}\right) - \Phi\left(\frac{12 - 13}{1.5}\right) \right] \\&= 1 - (.9082 - 0.2514) \\&= 0.3432\end{aligned}$$

Alternatively, with R we get  $P(X < 12 \text{ or } X > 15) = 0.3437038$

3. Consider the following sample of subjects with prostate cancer. They were followed over a period of 30 years. They are classified according to their usual fish consumption.

Fish Consumption	Frequency
Never/Seldom	14
Small part of diet	2420
Moderate Part	2978
Large Part	42

- (a) Is the variable categorical, ordinal, and/or numerical?  
(b) Produce a graphical summary of these data.  
(c) Is it reasonable to compute a median for these data? If so, compute it and interpret.  
(d) Give the modal class.

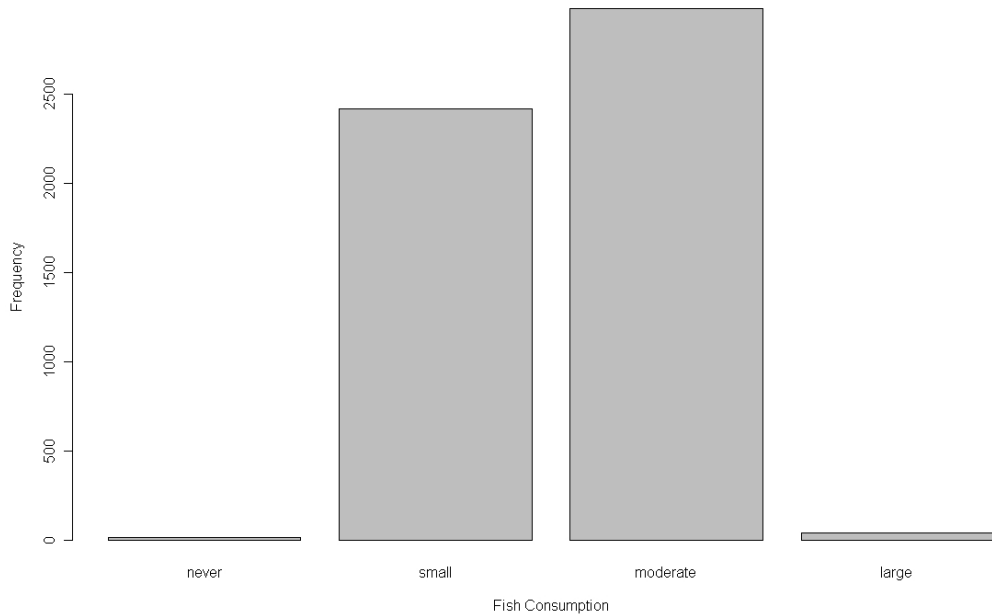
**Solution:**

- (a) The variable is categorical and ordinal. [We will also accept ordinal.]  
(b) We used the following R commands to produce the bar chart that is found below.

```

counts<-c(13,68,44,21,8)
names(counts)<-c("never","small","moderate","large")
barplot(counts,xlab="Fish Consumption",ylab="Frequency")

```



- (c) **[5 points]** Since it is an ordinal variable it is reasonable to give a median class, since we can interpret it. From the cumulative empirical distribution found below, we find the median. The median is moderate part. This means that about 50% of the subjects have a fish consumption that is moderate part or less or similarly we could say that 50% of the subjects have a fish consumption that is moderate part or more.

Fish Consumption	Frequency	cumulative rel. frequency
Never/Seldom	14	0.00257
Small part of diet	2420	0.44628
Moderate Part	2978	0.9923
Large Part	42	1.0
total	5454	

(d) The modal class is moderate part.

4. [5 points] Consider a study on acute otitis. Children were classified according the following variables. Below are values for one child in particular.

- age: 5 years
- sex: male
- profession of the father: skilled tradesman
- record of family on allergies (yes/no): yes
- parents use of tobacco (cig. per day) 0-9,10-19, 20 or more: 10 to 19
- number of otitis since the age of one: 2
- average hearing loss : 11.5 db
- articular pain (absent,moderate,intense): moderate

For each variable determine if it is categorical, ordinal and/or numerical. For numerical variables indicate if it is discrete or continuous.

**Solution:**

- age: numerical (continuous)
- sex: categorical
- profession of the father: categorical
- record of family on allergies (yes/no): categorical
- parents use of tobacco (cig. per day) 0-9,10-19, 20 or more: categorical, ordinal. **Note:** The categories do have a numerical interpretation so we would accept numerical (discrete) as well.
- number of otitis since the age of one: numerical (discrete)
- average hearing loss : numerical (continuous)
- articular pain (absent,moderate,intense): categorical and ordinal