# PubHlth 540 - Introductory Biostatistics
## Fall 2008
## Examination II
## SOLUTIONS

## 1. (10 points total)

a. **(3 points)**

How many ways can five photos be arranged on a shelf?

**Answer: 120**

Solution:

$5 \bullet 4 \bullet 3 \bullet 2 \bullet 1 = 5! = 120$. The solution is a permutation. You can think of putting five photos into 5 frames. First frame can have 5 different possible photos. Once a photo is placed in the first frame, we have 4 photos left. The second frame can only have 4 different photos, and so on. Therefore, we have $5 \bullet 4 \bullet 3 \bullet 2 \bullet 1$ possible arrangements.

b. **(3 points)**

You've registered for a class in world history. The course syllabus tells you that it will cover nine (9) books and that the final exam will cover five (5). How many different choices of five (5) books are possible?

**Answer: 126**

Solution:

The solution is a combinatorial.

$$\binom{9}{5} = \frac{9!}{(9-5)!5!} = 126$$

c. **(2 points)**

A student may receive a grade of EXCELLENT or GOOD or FAIR on exam I an a grade of PASS or FAIL on exam II. How many different outcomes of the form (EXAM I, EXAM II) are possible?

**Answer:** **6**

Solution:

total number of outcomes = (number of outcomes for exam I) • (number of outcomes for exam II)
total number of outcomes = 3 • 2 = 6

d. **(2 points)**

Consider your answer to problem #1C. What are these outcomes?

**Answer:**

| **Exam I** | **Exam II** | **Outcome** |
|---|---|---|
| Excellent | Pass | 1 |
| Excellent | Fail | 2 |
| Good | Pass | 3 |
| Good | Fail | 4 |
| Fair | Pass | 5 |
| Fair | Fail | 6 |

## 2. (5 points total)

Consider a binomial population for which N=6 and $\pi = 0.3$. What proportion of the population does X=2 represent?

**Answer: 32.41%**

Solution:

The solution is a binomial probability, the probability of observing two "successes" out of 6 trials, with the probability of each "success" being equal to 0.3.

$$P(X=2)=\binom{6}{2}(0.3)^2(1-0.3)^{6-2}=0.3241$$

**3. (25 points total)**

In 1889 a study was conducted by Geissler. Investigated were 53,680 families with 8 children. Of interest was an investigation of the probability of male gender. Is it .50? Does the probability of male gender change with birth order?

Geissler's observations are presented in the table below. As an example of its interpretation, the first row of this table is telling you that there were 215 families of 8 children in which the number of males was zero.

| # of male births among 8 children | # Families | % of Families |
|---|---|---|
| 0 | 215 | 0.4 |
| 1 | 1485 | 2.77 |
| 2 | 5331 | 9.93 |
| 3 | 10649 | 19.84 |
| 4 | 14959 | 27.87 |
| 5 | 11929 | 22.22 |
| 6 | 6678 | 12.44 |
| 7 | 2092 | 3.90 |
| 8 | 342 | 0.64 |
| **Total** | **53680** | **100%** |

a. **(5 points)**

Estimate π, the probability that a randomly selected birth is male. Assume that all births have the same probability of being male.

**Answer: 0.5147**

Solution:

The solution is the total number of male births / total number of births

$$\text{Total number of male births} = \sum_{i=1}^{8} (i = \#\text{of male births})_i (\#\text{families with } i \text{ males})_i = 221023$$

$= [\mathbf{0}](215) + [\mathbf{1}](1485) + [\mathbf{2}](5331) + \ldots + [\mathbf{8}](342)$

Total number of births = 8 • (# families) = 8 • 53680 = 429440

$\hat{\pi}$ = (221023) / (429440) = 0.5147

b. **(5 points)**

Using your answer for π in problem #3a, compute the binomial probabilities for number of boys in a family of eight equal to X = 0, 1, 2, 3, 4, 5, 6, 7, and 8. In reporting your answers, complete the following table.

| **# Male birth Outcomes X = x** | **Prob [ X = x ]** |
|---|---|
| 0 | **0.0031** |
| 1 | **0.0261** |
| 2 | **0.0969** |
| 3 | **0.2055** |
| 4 | **0.2725** |
| 5 | **0.2312** |
| 6 | **0.1226** |
| 7 | **0.0372** |
| 8 | **0.0049** |
| **Total** | **1** |

Solution:

Each Prob [ X = x ] is a binomial probability with number of trials N = 8 and π = .5147. These can be calculated using the applet given on http://onlinestatbook.com/analysis_lab/binomial_dist.html

$$P(X=0) = \binom{8}{0}(0.5147)^0(1-0.5147)^{8-0} = 0.0031$$

$$P(X=1) = \binom{8}{1}(0.5147)^1(1-0.5147)^{8-1} = 0.0261$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$P(X=8) = \binom{8}{8}(0.5147)^8(1-0.5147)^{8-8} = 0.0049$$

c. **(5 points)**

Using your answers to problem #3b, compare the observed **frequencies** of 0, 1, 2, 3, 4, 5, 6, 7, and 8 males to the expected **frequencies** of 0, 1, 2, 3, 4, 5, 6, 7, and 8 males. In obtaining the expected frequencies, assume that the binomial probability model is correct. In reporting your answer, complete the following table.

| # of male births among 8 children | Observed # Families | Expected # of Families |
|---|---|---|
| 0 | 215 | **166** |
| 1 | 1485 | **1401** |
| 2 | 5331 | **5202** |
| 3 | 10649 | **11031** |
| 4 | 14959 | **14628** |
| 5 | 11929 | **12411** |
| 6 | 6678 | **6581** |
| 7 | 2092 | **1997** |
| 8 | 342 | **263** |
| **Total** | **53680** | **53680** |

Solution:

Each expected # of families = Prob [ X = x ] • [Total # of families]
→
Expected # families with **0** male births = [ .0031 ] • [53680] = **166**
Expected # families with **1** male birth = [ .0261 ] • [53680] = **1401**
etc,

d. **(5 points)**

Now, using your answers to problem #3c, compare the observed **relative** frequencies of 0, 1, 2, 3, 4, 5, 6, 7, and 8 males to the expected **relative** frequencies. Again, assume that the binomial probability model is correct. In reporting your answer, complete the following table.

| **# of male births among 8 children** | **Observed proportion % Families** | **Expected proportion % of Families** |
|---|---|---|
| 0 | 0.4 | **. 31%** |
| 1 | 2.77 | **2.61%** |
| 2 | 9.93 | **9.69%** |
| 3 | 19.84 | **20.55%** |
| 4 | 27.87 | **27.25%** |
| 5 | 22.22 | **23.12%** |
| 6 | 12.44 | **12.26%** |
| 7 | 3.90 | **3.72%** |
| 8 | 0.64 | **0. 49%** |
| **Total** | **100%** | **100%** |

Solution:

Solution is a translation. The expected proportion of families with X=x males is the same as the binomial probability of X=x. Thus, the third column here is the same as the answer you produced for #3b.

e. **(5 points)**

What is your interpretation of Geissler's data with respect to the question of whether or not the probability of male birth is 0.50 and whether or not it is the same, regardless of birth order? Express your opinion in **1-3 sentences** written for a reader who is intelligent but who is not familiar with statistical jargon.

**Answer:**
**Comparison of the relative frequencies of each number of males among families with 8 children, observed versus expected, does not reveal any startling disparities. Thus, these data are consistent with the chances of male gender being constant and independent from birth to birth.**

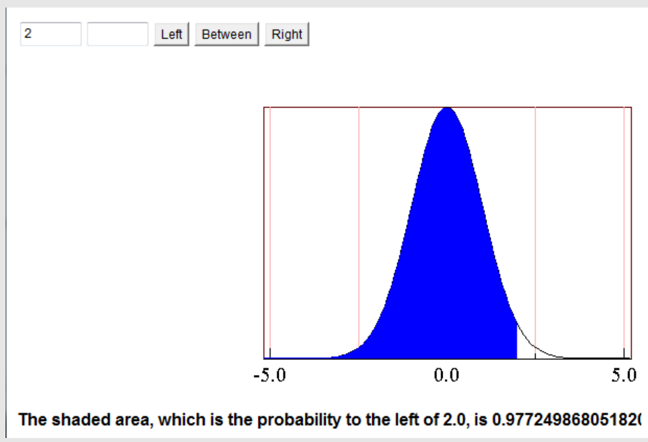
**4. (10 points total)**

A random variable Z is distributed standard normal, so that $\mu = 0$ and $\sigma = 1$.

a. **(2 points)**

Find Probability [ $Z \leq 2$ ]

**Answer: .9772**




b. **(2 points)**

Find Probability [ $Z \leq -1$ ]

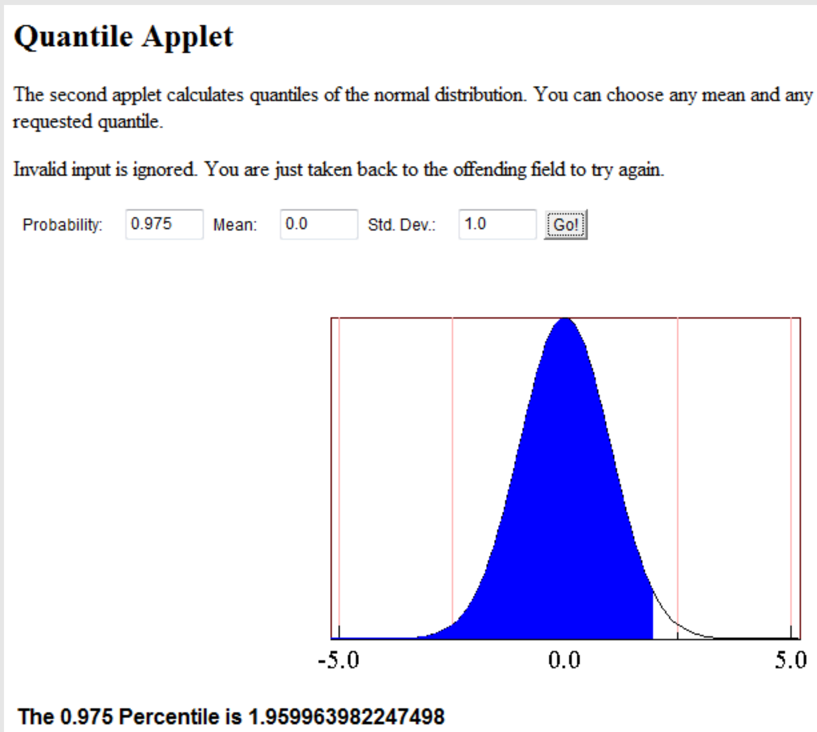
**Answer: 0.159**

Solution:
Same approach as that used for exercise #4a.

c. **(2 points)**

Find $Z_{.975}$, the value of the 97.5$^{th}$ percentile

**Answer: 1.96**



d. **(2 points)**

Find Probability [ $0.4 \leq Z \leq 1$ ]

**Answer: 0.186**

Solution:

The solution uses the same approach as that used for exercise #4a.

e. **(2 points)**

Find Probability [ $Z \leq -1.96$ or $Z \geq +1.96$ ]

**Answer: 0.05**

Solution:

P[ $Z \leq -1.96$ or $Z \geq +1.96$ ] = 1 - P[ $-1.96 \leq Z \leq +1.96$ ] = 1 - 0.950 = 0.05

**5. (10 points total)**

A random variable X is distributed normal with parameters $\mu = 2$ and $\sigma^2 = 4$.

a. **(2 points)**

Find Probability [ $X \le 0$ ]

**Answer: 0.159**

Solution:

Since $\mu \neq 0$ and $\sigma^2 \neq 1$ we cannot directly use the probability applet http://www-stat.stanford.edu/~naras/jsm/FindProbability.html.

Solution requires standardizing the desired probability calculation for the random variable X to an equivalent probability calculation for a random variable Z (also called "Z-score") that is distributed standard normal.

The transformation is given by $Z = \frac{X - \mu}{\sigma}$.

$P[ X < 0 ]$

$= P[ \frac{X-\mu}{\sigma} < \frac{0-\mu}{\sigma} ]$

. $= P[ Z < \frac{0-\mu}{\sigma} ]$ .

$= P[ Z < \frac{0-2}{\sqrt{4}} ]$

$= P[ Z < -1 ]$

Using the applet we find that $P[Z < -1] = 0.159$

b. **(2 points)**

Find Probability [ X $\geq$ -1 ]

**Answer: 0.933**

Solution:

$$P[X \geq -1]$$
$$= P\left[Z \geq \frac{-1-2}{\sqrt{4}}\right]$$
$$= P[Z \geq -1.5]$$
$$= 1 - P[Z \leq -1.5]$$
$$= 1 - 0.067$$
$$= 0.933$$

c. **(2 points)**

Find Probability [ 1 $\leq$ X $\leq$ 3 ]

**Answer: 0.383**

Solution:

$$P[1 \leq X \leq 3]$$
$$= P\left[\frac{1-2}{\sqrt{4}} \leq Z \leq \frac{3-2}{\sqrt{4}}\right]$$
$$= P[-0.5 \leq Z \leq 0.5]$$
$$= 0.383$$

d. **(2 points)**

Find Probability [ $X \leq 0.66$ or $X \geq 2.54$ ]

**Answer: 0.645**

Solution:

$$P[\ X\ < 0.66\ or\ X\ > 2.54\ ] = 1 - P[0.66 \leq X \leq 2.54] = 1 - P[\frac{0.66-2}{\sqrt{4}} \leq Z \leq \frac{2.54-2}{\sqrt{4}}] =$$

$$= 1 - P[-0.66 \leq Z \leq 0.27] = 0.645$$

e. **(2 points)**

Find $X_{.95}$, the value of the 95th percentile.

**Answer: 5.290**

Solution:
Working the standardization formula back to a solution for X, as on page 21 of the lecture notes,

$$X_{.95} = [\sigma]\ Z_{.95} + \mu$$

$$= [2](1.645) + 2$$

$$= 5.29$$

## 6. (10 points total)

Suppose it is known that a population distribution of a random variable X is normal with parameters $\mu = 1$ and $\sigma^2 = 9$. Simple random samples of sample size n=9 are taken and the sample means $\bar{X}_{n=9}$ are obtained.

(a) **(2 points)**

What is the correct probability model for the sampling distribution of $\bar{X}_{n=9}$? What are the values of its mean ($\mu$) and variance parameters ($\sigma^2$ )?

**Answer:** $\mu_{\bar{X}} = 1, \quad \sigma^2_{\bar{X}} = 1$

Solution:

$$\mu_{\bar{X}} = \mu = 1$$

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} = \frac{9}{9} = 1$$

(b) **(3 points)**

Find Probability [ $1 < \bar{X}_{n=9} \leq 2.85$ ]

**Answer: 0.4678**

Solution:

$$\Pr[\, 1 < \bar{X}_{n=9} < 2.85 \,]$$

$$= \Pr\left[\frac{1-1}{1} < \frac{\bar{X}_{n=9} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} < \frac{2.85-1}{1}\right]$$

$$= \Pr[\, 0 < \text{Z-score} < 1.85 \,]$$

$$= .4678$$

(c) **(5 points)**

Consider next a new random variable $W = 4\bar{X}_{n=9}$. What is the correct normal probability model for the sampling distribution of W? Specifically, what are the values of its mean ($\mu$) and variance parameters ($\sigma^2$)?

**Answer:** $\mu_W = 4, \quad \sigma^2_W = 16$

Solution:
Recall what happens when a random variable is multiplied by a constant:

$$\mu_W = E[W] = E[4\bar{X}_{n=9}] = (4)E[\bar{X}_{n=9}] = (4)\mu_{\bar{X}} = (4)[1] = 4$$

$$\sigma^2_W = Var[W] = Var[4\bar{X}_{n=9}] = (4^2)Var[\bar{X}_{n=9}] = (16)\sigma^2_{\bar{X}} = (16)[1] = 16$$

## 7. (10 points total)

Interestingly, it is generally agreed that we are poor probabilists. The citation source of this exercise will be given to you in the solutions. Here is the question

> *"A certain town is served by two hospitals. In the larger hospital, about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. However, the exact percentage varies from day to day. Sometimes it may be higher than 50%, sometimes lower. For a period of one year, each hospital recorded the days on which more than 60% of the babies born were boys. Which hospital do you thing recorded more such days? The larger hospital, the smaller hospital, or where they about the same (that is within 5% of each other)?*

In developing your answer to this question, you are being asked to work with two binomial probability distributions, once for the larger hospital and once for the smaller hospital. ***Hint: What is the expected value of each of the two binomial distributions and how do they compare?***

**Answer: The smaller hospital, with expected number of such days = 55 (versus 24)**
**When the number of births total is smaller, each individual birth represents a larger proportion of the total.**

Solution:

The solution involves the calculation of two binomial probabilities, one for the outcomes of one day at a given hospital and the other for the expected number of outcomes over an entire year.

Larger Hospital

Binomial Model for outcomes of one day:
\# trials, N = 45
$\pi = .5$
60% of 45 = 27
Event of "more than 60% boys" is $X > 27$
Pr [ $X > 27$ ] = **.0676**

Binomial Model for one year
\# trials, N = 365
$\pi = .0676$
Expected # days w more than 60% boys
= [ N ] [ $\pi$ ] = [365][.0676]
= **24.674**

Smaller Hospital

Binomial Model for outcomes of one day:
\# trials, N = 15
$\pi = .5$
60% of 15 = 9
Event of "more than 60% boys" is $X > 9$
Pr [ $X > 9$ ] = **.1509**

Binomial Model for one year
\# trials, N = 365
$\pi = .1509$
Expected # days w more than 60% boys
= [ N ] [$\pi$] = [365][.1509]
= **55.0785**

**8. (10 points total)**

The duration of gestation in healthy humans is approximately 280 days with a standard deviation of 10 days.

(a) **(5 points)**

Under the assumption of normality, what proportion of (healthy) pregnant women will be overdue by 2 weeks or more?

| |
|---|
| **Answer: 8.07%** |
| Solution:<br><br>Solution is a normal probability calculation for X distributed Normal with $\mu = 280$ and $\sigma = 10$.<br>"Overdue by 2 weeks or more" corresponds to $X \geq (280 + 14)$.<br>Thus, want Pr [ $X \geq 294$ ]<br><br>$\Pr[X \geq 294]$<br>$= \Pr\left[ Z \geq \frac{294 - 280}{10} \right]$<br>$= \Pr[Z \geq 1.4]$<br>$= 0.0807$ |

(b) **(3 points)**

Suppose now that, typically, there are 200 births per week at the UMass Memorial Hospital. How many of these births would you expect to be premature by 4 weeks or more?

**Answer: 0.52 births**

Solution:
"Premature by 4 weeks or more" corresponds to X $\leq$ (280 - 28).
Thus, want Pr [ X $\leq$ 252 ]

$$\Pr[X \leq 252]$$
$$= \Pr\left[ Z \leq \frac{252 - 280}{10} \right]$$
$$= \Pr\left[Z \leq -2.8\right]$$
$$= 0.0026$$

Thus, over 200 births, the expected number of premature babies is = [ 200 ] [ .0026 ] = **0.52**

(c) **(2 points)**

Using your answer to problem #8b, suppose that all premature babies are put into the neonatal intensive care unit. In **1-3 well written sentences**, what is your recommendation to the administration of UMass Memorial Hospital with respect to the size of the neonatal intensive care unit. ***Note – I am not looking for a particular answer here and I am not looking for formulas and calculations. I'm looking for you to show me that you have an understanding of the work you did in problem 8b.***

**Answer:**

**With 200 births per week, each with a .26% chance of being premature, the expected number of such babies per week is less than one. The expected number is 0.52. If length of stay in the neonatal intensive care is greater than 1 week, then the hospital can expect to require more than one such bed at any given time. A reasonable recommendation to the administration is to build neonatal intensive care unit capacity to accommodate at least 1-2 beds, with more depending on anticipated length of stay.**

**9. (10 points total)**

(a) (**5 points**).

A factory is producing thermometers and testing their quality with readings in a room that is known to be 32 degrees. Assume that the readings of a thermometer are distributed Normal with a mean of 32 degrees and a standard deviation of 2 degrees. If 3% of thermometers will be discarded because their readings are too low and another 3% of thermometers will be discarded because their readings are too high, find the two readings that separate the rejected thermometers from the accepted thermometers.

| |
|---|
| **Answer: 28.2 and 35.8** |
| Solution:<br><br>Solution requires obtaining the 3rd and 97th percentile of the Normal distribution with $\mu = 32$ and $\sigma = 2$<br><br>$X_{.03} = \sigma\,[Z_{.03}] + \mu = (2)[-1.8808] + 32 = 28.2384$<br>$X_{.97} = \sigma\,[Z_{.97}] + \mu = (2)[+1.8808] + 32 = 35.7616$ |

(b) **(5 points)**

The Air Force uses ACES-II ejection seats that are designed for men who weigh between 140 lb and 211 lb. Suppose it is known that women's weights are distributed Normal with mean 143 lb and standard deviation 29 lb. What proportion of women have weights that are ***outside*** the ACES-II ejection seat acceptable range?

**Answer: 46.83%**

Solution:
Solution is a normal probability calculation where $\mu = 143$ and $\sigma = 29$

$$1 - \Pr[\,140 < X < 211\,]$$

$$= 1 - \Pr\left[\frac{140-143}{29} < \frac{X - \mu}{\sigma} < \frac{211-143}{29}\right]$$

$$= 1 - \Pr\left[-0.1034 < \text{Z-score} < 2.3448\right]$$

$$= .4683$$