Midterm Exam Winter 2008 Answer sheet

a) ***What factors affect the chances of using health care?*** **Using table 1, briefly describe the relationship of healthcare use to other factors. [Use no more than about ½ page.]**

All three factors—income, health status, and health insurance—are associated with the use of health care, though the patterns differ by the type of care.

Higher **family incomes** are associated with the use of any health care and preventative health care, but the chance of a hospital stay is lower for those with higher incomes.

Women with worse **health status** (poor or fair) were more likely to use at least some health care and much more likely to have a hospital stay (22% versus 9 percent). Preventative health care, however, was more likely among those in better health.

Those with no **health insurance** were much less likely to use any health care or preventative health care than were those with either private or public insurance, but rates for hospital stays for the uninsured were similar to rates for those with private insurance. Hospital visits were most likely for those with public insurance (perhaps because of the higher rates for elderly with Medicare).

b) ***How many people use health care?*** **Find the overall percentage of women who used at least some health care. Also find the percentage of those with some preventative care within the last two years.**

To find the overall percentage with at least some health care we need to compile the conditional rates we have for a set of categories (say health status). Here HC=any health care and FAIR and GOOD are the health statuses.

P(HC) =P(HC and FAIR) + P (HC and GOOD
=P(HC|FAIR) x P(FAIR) + P(HC|GOOD) x P(GOOD)
=(.96)(.14) + (.90)*(.86)
=.91
P(PREV) =P(PREV and FAIR) + P (PREV and GOOD
=P(PREV|FAIR) x P(FAIR) + P(PREV|GOOD) x P(GOOD)
=(.90)(.14) + (.94)*(.86)
=.94
Note that the large group of people with good or better health really carry more weight in the overall mean.

So 91 percent of all women had some health care in the last year and 94 percent had had some preventative care in the last 2 years. Note the preventative care measure covers a longer period—that's the only reason it can be higher than the percentage for any health care.

c) How many people seem disadvantaged in health care? What percent of people overall were both poor/near poor and did not receive any health care in the last year?

Here we want Pr(Poor and no HC). To get this we will need the multiplicative formula and the complement. Here Poor is the lowest income level and HC is any health care.

We have Pr(Poor)= .16 and Pr(HC|Poor)=.88

The complement gives us: Pr(no HC|Poor)= 1-Pr(HC|Poor)=1-.88=.12

Now we can use the multiplicative to get

Pr(Poor and no HC)=Pr(no HC|Poor)*Pr(Poor)=.12*.16**=.02**

Also, what percent were uninsured and received no health care?

Similarly we want Pr(Uninsured and no HC):

We have Pr(Uninsured)= .11 and Pr(HC|Uninsured)=.71

The complement gives us:
Pr(no HC|Uninsured)= 1-Pr(HC|Uninsured)=1-.71=.29

Now we can use the multiplicative to get :

Pr(Uninsured and no HC)=Pr(no HC|Uninsured)*P(Uninsured)=.11*.29**=.031**

Because the group size for uninsured (11 percent) is low, this overall probability of no HC isn't much more than that for the poor group even though the proportion of the uninsured with no health care is more than double (29 percent versus 12 percent).

So 2 percent of all women were poor and had no health care and 3.1 percent were uninsured and had no health care. Of course, these two groups are likely to overlap!

d) *How does insurance affect health care use?* Of those who had no health care in the last year, what was the probability of being uninsured? Similarly, what percentage of those with no preventative care had no insurance? What was the percentage for those with no hospital stays?

Here we need Pr(Uninsured | no HC)

We can use Baye's theorem or work from the pieces of the conditional probability formula:
Pr (Uninsured | no HC) = Pr (Uninsured and no HC)/ Pr (no HC)

In part c, we calculated the Pr(Uninsured and no HC)= .031 and in part a we calculated Pr (HC)= .91 so Pr (no HC) = 1-.91 = .09

So Pr (Uninsured | no HC) = .031/.09 = **.34**

For preventative care, we need Pr(Uninsured | no Prev) :
Pr (Uninsured | no Prev) = Pr (Uninsured and no Prev)/ Pr (no Prev)

In part a we calculated Pr (Prev)= .94 so Pr (no Prev) = 1-.94 = .06

Pr(Uninsured and no Prev)=Pr (no Prev|uninsured)*Pr(Uninsured)
= (1-.80)*(.11)=.021
So Pr (Uninsured | no Prev) = .021/.06 = **.325**

For hospital stays, we need Pr(Uninsured | no HOS) :

Pr (Uninsured | no HOS) = Pr (Uninsured and no HOS)/ Pr (no HOS)

Calculate Pr (no HOS)= .**89** as in part a (or see spreadsheet).

Pr(Uninsured and no HOS)=Pr (no HOS | uninsured)*Pr(Uninsured)
= (1-.06)*(.11) =.10
So Pr (Uninsured | no HOS) = .10/.89 = **.112**

So the uninsured make up 34 percent of those with no health care, 32 percent of those with no preventative care, but only 11.2 percent of those with no hospital stays. Note that they are overrepresented among those with no preventative or other care, but fairly representative among those with no hospital stays.

e) ***What is the average level of health status?*** **Use Table 2 to calculate the mean level of health status (on the four point scale) for each category of insurance (separately). Also calculate the standard deviation for each group.**

To get the mean level for one group you need to sum up the values weighted by the probability of each value. So for the private insurance group:

$$\mu = E(X) = \sum X\,P(X) = (1)(.04) + (2)(.06) + (3)(.62) + (4)(.28)$$
$$= \mathbf{3.14}$$

Similar calculations give **2.80** and **3.01** for the public and no insurance groups, respectively.

To get the variance, we need to get the distance from the mean for each value, square that distance, weight it by its probability and add them up. So for the private insurance group:

$\sigma^2 = E(X-\mu)^2 = \sum (X-\mu)^2 P(X) =$
$(1\text{-}3.14)^2(.04) + (2\text{-}3.14)^2(.06) + (3\text{-}3.14)^2(.62) + (4\text{-}3.14)^2(.28) = .49$

The standard deviation is then:
$\sigma = \sqrt{\sigma^2}$ **= .71**

The standard deviations for the public insurance and no insurance groups are **.87** and **.81**—larger than that of the private insurance group.

f) ***What might a new insurance program accomplish?*** **Suppose there was a new public insurance program targeting those who were poor or near poor and uninsured. Use the information from Table 1 or your previous calculations to estimate the percentage of the population the program might serve.**
Let's try to get some high and low boundaries around how many people might be Poor/Near poor and uninsured. P(Poor and Uninsured) [I'll use Poor for Poor or near poor here.]

First, the probability cannot be more than Pr(Poor)=.16 or than Pr(Uninsured)=.11, so the **very upper bound is 11 percent**. [*Be sure you understand why the intersection cannot be more than either component*!]

Second, if being uninsured and poor were statistically independent, then the proportion of the uninsured would be the same among the poor as among the other groups. [In real life, of course, the rate of being uninsured is MUCH higher among low income families.] A low bound estimate, assumes independence:

Pr(Uninsured and Poor)=Pr(Uninsured|Poor)*Pr(Poor)
IF independent: then Pr(Uninsured| Poor)= Pr(Uninsured)=.11
Pr(Uninsured and Poor)=Pr(Uninsured)*Pr(Poor)=.16*.11=.017

So a good low bound estimate is 1.7 %

To get a different estimate, I used information on the use of health care because we have information on the use of any health care use by both insurance status and income category.

Here's my strategy: I know Pr(Uninsured| no HC)=.34 (from part d) and can get Pr (Uninsured| HC)= .08 (using a similar strategy).

Since I know for the poor Pr(HC)=.88 and Pr (no HC)=1-.88=.12 (from Table 1) I can use these as weights to get an estimate of

Pr (Poor and Uninsured)=Pr(Poor and Uninsured and no HC) + Pr(Poor and Uninsured and HC)
So Pr(Poor and Uninsured)= .88*(.34) + .12(.08)=.11 or **11 percent**

Because this estimate is the same as my upper bound, I'm inclined to believe that the real estimate of the number of people who are uninsured and poor or near poor is very close to 11 percent.

g) Write a short description of your results for your committee (less than 1 page typed). They will not look at the rest of your exam so be sure that you describe your results for them in the context of the issues they are interested in. Discuss the policy implications and any caveats you have about the data or information.

A recent study of healthcare for American women showed that use of health care is related to the availability of insurance, income, and health status. Overall use of health care was high—91 percent of women reported using health care in the last year. But use of care was lower for those with no insurance and lower incomes.

Overall, 9 percent of women reported using no health care in the last year. Although good health could explain some of this, rates of no care were much higher for those with no insurance (29 percent) and somewhat higher for the poor and near poor (12 percent) suggesting that lack of insurance and income play a role. Preventative care in the last two years was also less likely for those without insurance (20 percent) and low income (12 percent) than for the general population (6 percent).

Two percent of the population was poor and had no health care in the last year and 3 percent were uninsured and had no health care—small percentages, but significant numbers of people for the nation.

About a third of those with no health care had no insurance and a similar percentage of those with no preventative care were uninsured. Conversely, hospital stays were most common for those with public insurance (perhaps attributable to medicare for the elderly), so only 11 percent of those without hospital stays were uninsured.

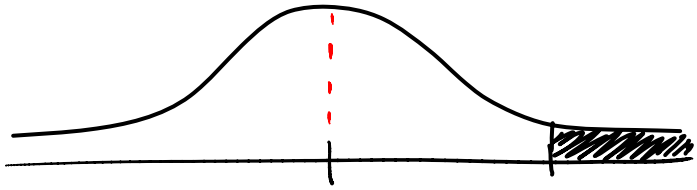
Most people report being in good or excellent health (86 percent), though the percentage was higher for those with private health insurance (90 percent) than for those with no insurance (80 percent) or public insurance (70 percent).

A new public insurance program could target the 2 to 11 percent of people who are very low income and currently uninsured. A new program has potential to increase access to healthcare for those without insurance, but the design of the program is critical if we want to avoid drawing people away from other sources of insurance such as employers or other public federal programs.

These data provide important information, but we should triangulate these conclusions with additional information sources. Patterns could be different for men (not addressed by the study) because their differing patterns of employment and child-rearing could affect their access to insurance and use of health care could differ. In addition, further information on the characteristics of those without insurance could help explain whether adding a new program would increase access to needed health care.

Question II

1. First, draw the graph of the area that you are interested in. If X is the number of minutes spent watching TV per day, we want to know P(X>180).



Next, convert your distribution into the standard normal by converting the value we are interested in (X=180 minutes) into a z-score.

$$\frac{X-\mu}{\sigma}=\frac{(180-90)}{86}=1.05=Z$$

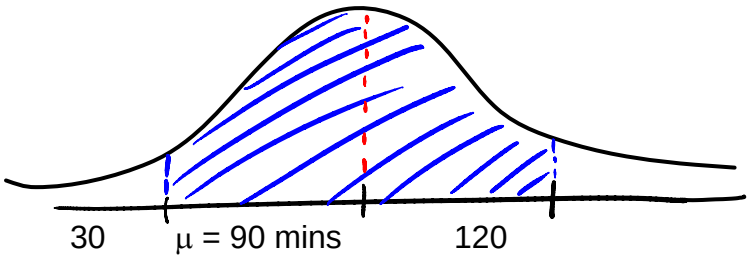
Now, find the area under the standard normal curve for (Z > 1.05). If you use the standard normal table in your text and look up the value for Z=1.05, you get a probability of 0.3531. But this is the area between the mean and Z = 1.05 (or the P(0 < Z < 1.05), but we want the tail. The table gives us probabilities under half of the standard normal, so to get the tail:

P(Z > 1.08) = 0.5 – 0.3531 = **0.147** = P(X > 180)

You could also have gotten the probability density from the online calculators or Excel.

We expect that 14.7% of 6-8 years old watch more than three hours of TV per day.

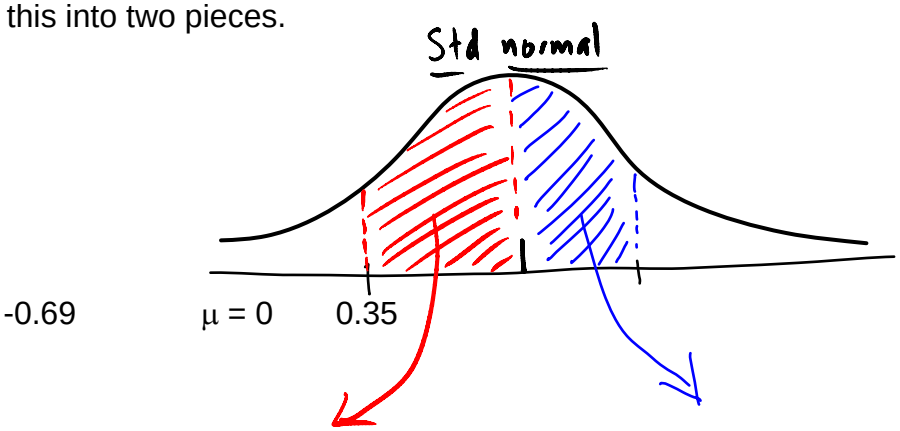
2. Start by drawing your graph. We want to find P(30 < X < 120).



Now we translate these into the standard normal. Here we have two values that we are interested in.

$$\frac{X_{upper}-\mu}{\sigma}=\frac{(120-90)}{86}=0.35=Z_{upper} \qquad \frac{X_{lower}-\mu}{\sigma}=\frac{(30-90)}{86}=-0.69=Z_{lower}$$

So now we want to find P(-0.69 < Z < 0.35) for the standard normal distribution. Let's split this into two pieces.



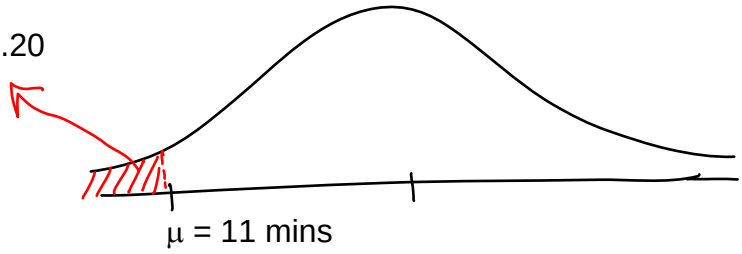
(-0.69 < Z <0) = P(0 > Z > 0.69) because ormal curve is symmetric, so you can look he area for Z=0.69 in the standard normal able, which is **0.2549**.	(0 > Z > 0.35) ou can look this up in the standard norma able, which is **0.1368**.

Now add the two probabilities, so:
P(-0.69 < Z < 0.35) = 0.2549 + 0.1368 = **0.39** = P(30 < X < 120).

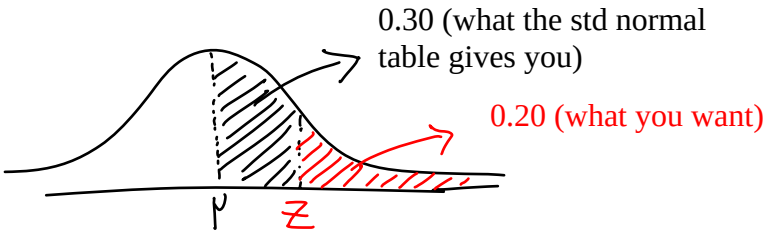
We expect that 39% of 6-8yr olds watch TV between 30 minutes and two hours per day.

3. In this problem you are given the probability (20%), and you need to calculate the test score which corresponds to that probability. Let's define X this time as the number of minutes spent <u>reading</u> per day. Again, start with a picture:

P(X < ?) = 0.20



If you find the Z-score using the standard normal table, you need to find the Z corresponding to a density of 0.30.



That is Z= 0.84, or Z=-0.84 for our problem.

Now use the formula for z-scores to translate Z into X (reading scores)

$$z = \frac{X - \mu}{\sigma} \qquad \text{so} \qquad -0.84 = \frac{X - 11}{20}$$

$$11 - 0.84 * 20 = X$$
$$-5.8 = X$$

So the cutoff (to identify the 20% of children who read the least) should be those who read -5.8 minutes of reading per day or less, but wait! Reading minutes can't be less than 0, so all these must be piled up at 0. We could further find out how many kids total are piled at 0 by looking up the probability Z score for X=0.

$z = \frac{X - \mu}{\sigma}$ so $-0.55 = \frac{0 - 11}{20}$ and P(Z<-.55)=.2912.

So 29 percent of kids had 0 minutes of reading (and would have read less if it were possible!)