

Introductory Biostatistics 1
Homework 3
Due: September 24, 2020

1.

A psychiatric researcher wishes to evaluate a new diagnostic instrument for psychotic illness. The test is comprised of 20 questions regarding individual psychiatric symptoms. A positive test is one in which a score of at least 16 of the 20 symptoms. The sensitivity of the test was determined to be 0.96, while the specificity is 0.99. The following gives the psychotic disease prevalence rates for two subpopulations of people:

Subpopulation	Prevalence
General Population	0.001
Homeless	0.1

- a. Compute predictive values of a positive and negative test for each of the above subpopulations.
- b. Suppose that predictive value of a positive test of greater than 0.5 is deemed acceptable. Under this criterion, what subpopulations should receive the test.
- c. Suppose that we have a sample of 1000 homeless with positive test results.
 - i. How many of those people do you expect to actually have psychotic illness?
 - ii. How many of those people do you expect not to have psychotic illness?

2.

Take a look at the following article on diagnostic testing:

Gibot, S., et al. (2004). Soluble triggering receptor expressed on myeloid cells and the diagnosis of pneumonia. *New England Journal of Medicine* 150, 451-458.

- a. What disease are the authors attempting to diagnose? Be specific.
- b. What biomarkers are being considered for the disease?
- c. Figure 2 gives the ROC for the three biomarkers.
 - i. Give the areas under the curve for the three biomarkers. (Hint: look at the figure legend).
 - ii. Based on the ROC curves, which biomarker performed best? Which performed the worst?
- d. Which biomarker was selected by the authors? Was this the best performing biomarker? (Hint: see text pp. 455-456)
- e. What screening value was selected for the selected biomarker?
- f. What is the sensitivity and specificity of the biomarker under the selected screening value?

3.

Located in your neck, the thyroid gland produces Thyroxine, a hormone that regulates the conversion of oxygen and calories into energy. Hypothyroidism is a disease, resulting in the reduced production of thyroid hormones. Consider the use of thyroxine as a marker for hypothyroidism. Estimates of the sensitivities and specificities for various screening values are presented in the following table:

Screening Level	Sensitivity	Specificity	PV^+	PV^-
<1	0.0625	1.0000	1.00000	0.95298
<2	0.1562	1.0000	1.00000	0.95748
<3	0.1875	1.0000	1.00000	0.95899
<4	0.4375	1.0000	1.00000	0.97125
<5	0.5625	0.9892	0.73271	0.97725
<6	0.6875	0.9247	0.32457	0.98252
<7	0.7812	0.8065	0.17525	0.98592
<8	0.8750	0.6022	0.10376	0.98919
<9	0.9062	0.4194	0.07591	0.98837
<10	0.9062	0.2043	0.05655	0.97641
<11	0.9688	0.0860	0.05284	0.98126
<12	1.0000	0.0000		

The values of sensitivity and specificity appear in the STATA worksheet Thyroxine.dta. Note that in this case, low levels of thyroxine are used to screen for hypothyroidism.

- Construct the ROC curve for these data.
- Compute the area under the ROC curve.
- Does thyroxine appear to be a good marker for hypothyroidism? Justify your answer.
- In the above table, the predictive values of positive and negative tests are given for a population of people visiting a medical clinic for exhaustion, where prevalence of hypothyroidism is 0.05. Based on this table, what screening level should be used? Assume that false positives and false negatives are equally costly.