What is the sensitivity and specificity of a pneumonia model that always outputs positive?  In other words, the models says that every patient has the disease.

**1 point**

○ sensitivity = 0.5,  specificity = 0.5

○ sensitivity = 0.0, specificity = 1.0

○ sensitivity  = 1.0, specificity = 1.0

○ sensitivity = 1.0, specificity = 0.0

2. In some studies, you may have to compute the Positive predictive value (PPV) from the sensitivity, specificity and prevalence.

**1 point**

Given a sensitivity = 0.9, specificity = 0.8, and prevalence = 0.2, what is the PPV (positive predictive value)?

HINT: please check the reading item "Calculating PPV in terms of sensitivity, specificity and prevalence"

○ 0.18

○ 0.02

○ 0.9

○ 0.53

3. If sensitivity = 0.9, specificity = 0.8, and prevalence = 0.2, then what is the accuracy?

**1 point**

Hint: You can watch the video "Sensitivity, Specificity and Prevalence" to find the equation.

○ 0.75

○ 0.44

○ 0.82

○ 0.52

**4.** What is the sensitivity and specificity of a model which randomly assigns a score between 0 and 1 to each example (with equal probability) if we use a threshold of 0.7?

**1 point**

○ Not enough information to answer the question.

○ Sensitivity = 0.5, Specificity = 0.5

○ Sensitivity = 0.3, Specificity = 0.7

○ Sensitivity = 0.7, Specificity = 0.3

**5.** What is the PPV and sensitivity associated with the following confusion matrix?

**1 point**

Recall that

$$PPV = \frac{TruePositives}{positive\ predictions}$$

*Sensitivity* = How many actual positives are predicted positive?

|  | Test Positive | Test Negative |
|---|---|---|
| **Disease Positive** | 30 | 20 |
| **Disease Negative** | 70 | 10 |

○ PPV = 0.4, Sensitivity = 0.2

○ PPV = 0.6, Sensitivity = 0.33

○ PPV = 0.3,  Sensitivity = 0.6

○ Not enough information is given

**6.** You have a model such that the lowest score for a positive example is higher than the maximum score for a negative example. What is its ROC AUC?

**1 point**

HINT 1: watch the video "Varying the threshold".

HINT 2: draw a number line and choose values for the score that is the lowest prediction for any positive example, and choose another number that is the score for the highest prediction for any negative example. Draw a few circles for "positive" examples and a few "x" for the negative examples. What do you notice about the model's ability to identify positive and negative examples?

○ 1.0

○ Not enough information is given

○ 0.52

○ 0.82

7. For every specificity, as we vary the threshold, the sensitivity of model 1 is at least as high as model 2. Which of the following must be true?

**1 point**

○ The accuracy of model 2 is higher than model 1

○ The ROC of model 1 is at least as high as model 2

○ None of the above

○ The ROC of model 2 is higher than model 1

8. You want to measure the proportion of people with high blood pressure in a population. You sample 1000 people and find that 55% have high blood pressure with a 90% confidence interval of (50%, 60%). What is the correct interpretation of this result?

**1 point**

HINT: Please watch the video "Confidence interval" to help you answer this question.

○ There is a 5% chance that the true mean is less than 50%

○ If you repeated this sampling, the true proportion would be in the confidence interval about 90% of the time

○ With 90% probability, the proportion of people with high blood pressure is between 50% and 60%

○ If we repeated this sampling, the middle of the confidence interval would be 55%, 90% of the time

**9.** One experiment calculates a confidence interval using 1000 samples, and the another computes it using 10000 samples. Which interval do you expect to be tighter (assume they use the normal approximation)?

**1 point**

○ Not enough information

○ Cannot say with confidence

○ 1,000 samples

○ 10,000 samples