Congratulations! You passed!

Grade received 100% **To pass** 80% or higher

Go to next item

1. Which of the following is **not** an example of a clinical application of a prognostic model? 1 / 1 point

- ○ Determining who should receive drugs for reducing heart attack risk
- ◉ Detecting atrial fibrillation automatically using a EKG
- ○ Informing patients about their risk of developing illness
- ○ Determining who should receive end of life care

Correct
Correct answer. Prognosis involves predicting the risk of future events. Detecting atrial fibrillation involves determining something that has already happened, so it is NOT an example applying prognosis.

2. Recall the MELD score from the lesson. What is the output for a person with 1 / 1 point

Creatinine = 0.8 mg/dL,

Bilirubin total = 1.5 mg/dL,

INR = 1.3

Remember that the final score is multiplied by 10.

Please use natural logarithm instead of base 10 log.

You can also watch the video "Liver Disease Mortality" to review the calculation of the MELD score.

Variable	Coefficient
Ln Creatinine (mg/dL)	0.957
Ln Bilirubin total (mg/dL)	0.378
Ln INR	1.120
Intercept	0.643

- ◉ 8.76
- ○ 0.876
- ○ None of the above
- ○ 7.44

Correct
Evaluating the formula, we get ln(0.8)*0.957 + ln(1.5)*0.378 + ln(1.3)*1.12 + 0.643 = 0.876. Multiplying this by 10 we get 8.76.

3. You've fit a linear model with no interaction terms, and which include Age (in years) as an input feature of the model. Also, you don't multiply the sum product by any scaling number (unlike the MELD score, for instance). 1 / 1 point

The risk score for a patient measured today is 0.56.

The model's coefficient for age is 0.24.

What will this patient's risk score be one year later, if all other features remain the same?

- ○ 0.56
- ◉ 0.80
- ○ 0.24
- ○ Not enough information

Correct
Because the model is linear and since there are no interaction terms, and since it's given that the model does not multiply the sum product by a scaling factor, you have enough information to calculate the patient's risk score.

The only change in the features will be the age, which will increase by one year.

This will add 1×0.24 to the original risk score.

The new risk score will be $0.56 + 0.24 = 0.80$.

4. A linear risk model for the risk of heart attack has three inputs: Age, Systolic Blood Pressure (BP), and the interaction term between Age and Systolic Blood Pressure. The coefficients for Age, BP, and the interaction term are 0.1, 0.3, and 0.5. 1 / 1 point

Can you determine how an increase in blood pressure is affected by an increase in age?

HINT: here is the formula for the model:

$$y = (\beta_A \times Age) + (\beta_B \times BP) + (\beta_{AB} \times Age \times BP)$$

- ○ The effect of blood pressure on risk is independent of age
- ○ As you get older, the same increase in blood pressure leads to a SMALLER change in your risk of heart attack.
- ○ None of the above
- ◉ As you get older, the same increase in blood pressure leads to a LARGER change in your risk of heart attack.

Correct
Correct. If you factor out the BP feature from the given formula, you'll get:

$$y = (\beta_A \times Age) + ((\beta_B) + (\beta_{AB} \times Age)) \times BP$$

So the total coefficient of BP is

$$((\beta_B) + (\beta_{AB} \times Age))$$

Therefore at a higher Age, the total coefficient on BP is larger, so the same increase in BP also increases the predicted risk more than when the Age is lower.

5. If a feature x has range 0 to ∞, then what is the range of $\ln(x)$? 1 / 1 point

- (0, infinity)
- (-infinity, 1)
- (-infinity, infinity)
- None of the above

Correct
Recall that as $x \to 0$, then $ln(x) \to -\infty$,

and as $x \to \infty$ then $ln(x) \to \infty$.

Therefore after we log transform the variable, the range is $(-\infty, \infty)$.

Applying the natural log is helpful because it reduces the skew in the distribution of a feature covariate when the feature's range of values is strictly positive.

This is especially valuable in a linear regression, since the model treats positive changes the same a negative changes.

6. True or False: If a > b, then ln(a) > ln(b). 1 / 1 point

- False
- True

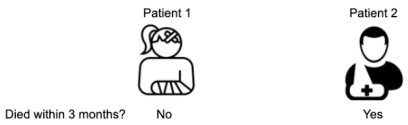
Correct
It helps to see a graph of $ln(x)$, since it is always increasing as x increases.

This means that the function is monotonic, and means that natural log maintains the order of the inputs.

This means that if $a > b$, then $ln(a) > ln(b)$.

This makes it a very reasonable transformation to apply, since it will preserve the order to the values in your dataset.

7. Which assignment of risk would make the following pair **concordant**? 1 / 1 point



- (0.5, 0.83)
- None of the above
- (0.44, 0.44)
- (0.76, 0.34)

Correct
Patient 2 had a worse outcome (died within 3 months) than patient 1.

For the pair to be concordant, Patient 2 should be assigned a higher risk score than patient 1.

8. What is the C-index for the following set of predictions? 1 / 1 point

Patient	Event	Risk
1	Yes	0.74
2	Yes	0.52
3	No	0.60
4	No	0.28

- 0.5
- 0.75
- 1.0
- 0.25

Correct
There are 4 permissible pairs ((1, 3), (1, 4), (2, 3), (2, 4)). Of these, only one is not concordant ((2, 3)), since 2 has a worse outcome but 3 has a higher risk score. Therefore ¾ of the permissible pairs are concordant, so since there are no ties this means the C-index is 0.75.

9. What is the C-index for a model which always outputs 0.6 for any patient regardless of their health outcome? 1 / 1 point

- 0.6
- There is not enough information to say
- 0.5
- 0.0

Correct
Because the model only outputs the same value for any patient, every permissible pair is going to be a risk tie.

Therefore, in the numerator of the c-index formula, every pair will get a weight of 0.5, so the overall c-index will be 0.5.

10. Model 1 has a c-index of 0.7 and Model 2 has a c-index of 0.6. Which is more accurate using a threshold of 0.5 for the risk score? **1 / 1 point**

In other words, if the risk score is 0.5 or higher, predict that the patient will have the disease in the future. If the risk score is < 0.5, predict that the patient will not have the disease.

- (●) There is not enough information to say
- ( ) Model 2
- ( ) They are equally as accurate
- ( ) Model 1

Correct

Like ROC, the c-index aggregates performance across all operating points (all thresholds).

It does not say anything about a particular threshold. A model may have a c-index of 1, but still have all the risk scores be above 0.5, and therefore have awful accuracy at that threshold (because all of its predictions would then be positive for the disease).

Therefore, the c-index does not say which model is more accurate if the threshold for the risk score is 0.5 (or any other value for the threshold).