1. **Imagine that you are collecting variables while participants attempted to shoot a soccer ball. Which of the following collected variables could be predicted using a logistic regression model?**

☐ Age (years)

☑ Whether a shot on goal traveled more than 20 feet

> ✓ Correct

☑ Scoring a soccer goal on a given shot
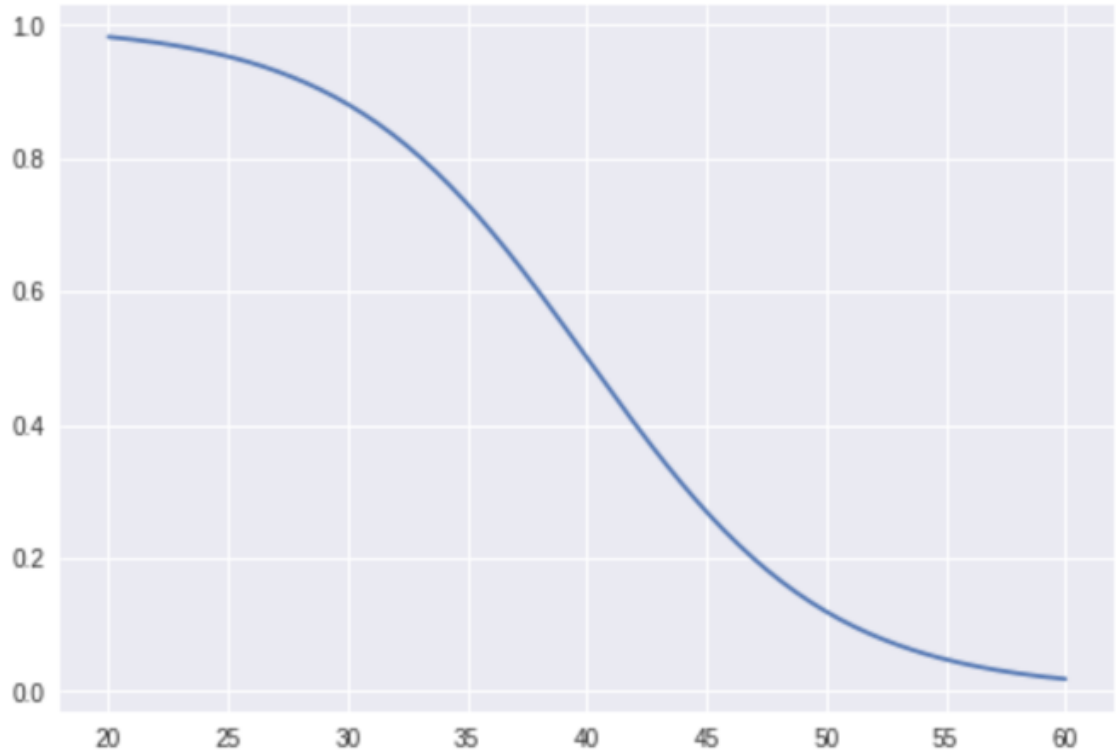
> ✓ Correct
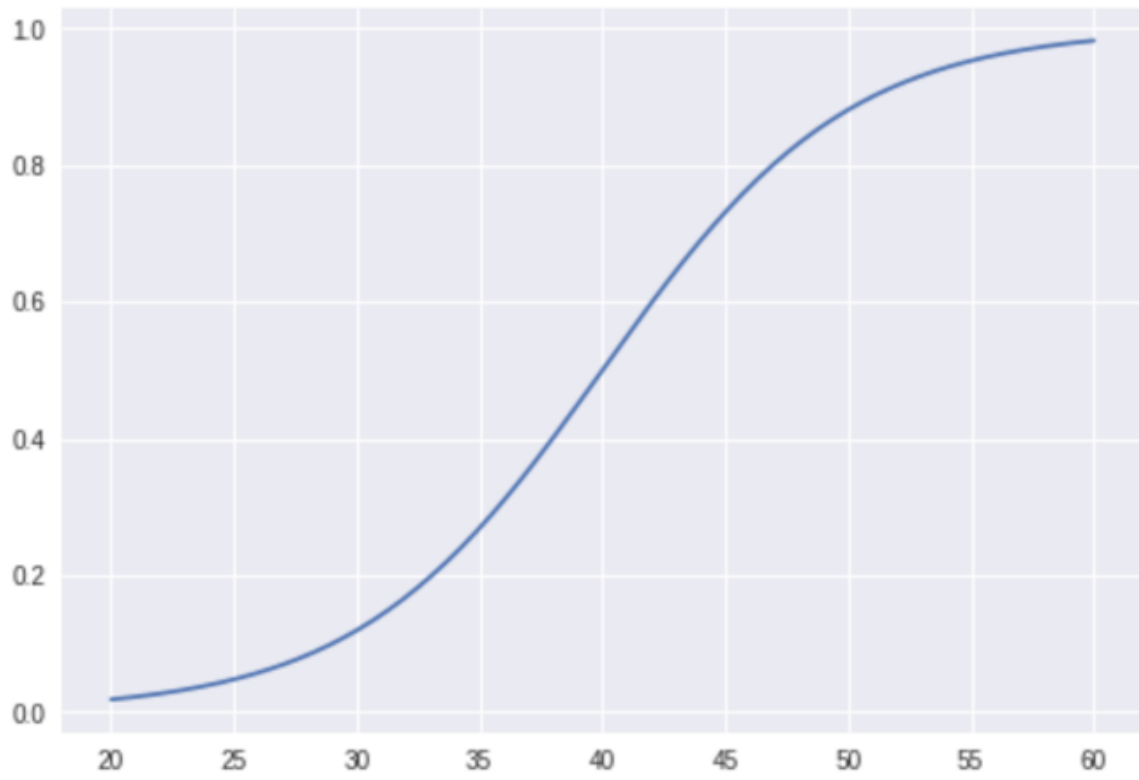
☐ Height

☑ Sex (male vs. female)

> ✓ Correct

2. **Which of the following is a possible form/shape for a logistic regression model, where the y-axis represents the probability of success?**

☑ Graph:

| ✓ Correct |

Graph:

✓ **Correct**

3. **Two probabilities have been transformed using the logit function. The two values after transformation are -2 and 0.25. Which of the two values corresponds to a higher original probability?**

○ -2

◉ 0.25

○ They are the same

○ Can't tell

✓ Correct

4. **NHANES records whether an individual has smoked 100 cigarettes or more. The next few questions will focus on fitting models to predict whether someone has smoked 100+ cigarettes.**

First, a model is fit using body mass index (BMI) as the variable to predict smoking status. The output is here:

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-0.4219	0.123	-3.426	0.001	-0.663	-0.181
BMI	0.0037	0.004	0.911	0.362	-0.004	0.012

What does the coefficient of 0.0037 mean?

○ For each increase by one in BMI, the probability of smoking 100 cigarettes increases by about 0.0037, on average.

○ For each increase by one in BMI, the odds of smoking 100 cigarettes increases by about 0.0037, on average.

◉ For each increase in one in BMI, the log odds of smoking 100 cigarettes increases by about 0.0037, on average.

○ For each increase in one in BMI, the odds of smoking 100 cigarettes increases multiplicatively by about 0.0037, on average.

✓ Correct

5. Next, a model is fit adding Age as an additional covariate to BMI as the variables predicting smoking status. The output is here:

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-1.2435	0.149	-8.366	0.000	-1.535	-0.952
BMI	0.0030	0.004	0.718	0.472	-0.005	0.011
Age	0.0169	0.002	10.349	0.000	0.014	0.020

What does the coefficient of 0.0169 mean in context?

○ For each increase of one in BMI, the odds of smoking 100 cigarettes increases by about 0.0169, on average.

○ For each increase of one in Age, the odds of smoking 100 cigarettes increases by about 0.0169, on average.

○ For each increase of one in Age, the log odds of smoking 100 cigarettes increases by about 0.0169, on average.

◉ For each increase of one in Age, the log odds of smoking 100 cigarettes increases by about 0.0169 while holding BMI constant, on average.

✓ Correct

6. **Based on the logistic regression with both Age and BMI as covariates, are the coefficients statistically significant at a two-sided 10% significance level?**

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-1.2435	0.149	-8.366	0.000	-1.535	-0.952
BMI	0.0030	0.004	0.718	0.472	-0.005	0.011
Age	0.0169	0.002	10.349	0.000	0.014	0.020

○ Both coefficients are significant

○ Neither coefficient is significant

○ Only the coefficient for BMI is significant

◉ Only the coefficient for Age is significant

✓ Correct

7. **The 95% confidence interval for the coefficient for Age is given above as (0.014, 0.020). If instead we wanted a 90% confidence interval, how would the width of the interval change?**

○ It would be wider

◉ It would be narrower

○ It would stay the same

○ Can't tell

✓ Correct

8. We'd like to predict the log odds of smoking 100+ cigarettes for a given individual using the logistic regression model with the two variables: BMI and Age. For an individual with a BMI of 22 who is 45 years old, what would the predicted log odds be?

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-1.2435	0.149	-8.366	0.000	-1.535	-0.952
BMI	0.0030	0.004	0.718	0.472	-0.005	0.011
Age	0.0169	0.002	10.349	0.000	0.014	0.020

- ● -0.417
- ○ 0.8265
- ○ 0.327
- ○ -0.7367
- ○ Can't tell

✓ Correct

9. The sample of adults surveyed in NHANES contains adults age 20-80 with BMIs of 14.5-64.6. For the individual with a BMI of 22 who is 45 years old, do you trust the predicted log odds calculated above as being reasonable?

- ○ No, this is extrapolation
- ○ No, this is interpolation
- ○ Yes, this is extrapolation
- ● Yes, this is interpolation

✓ Correct

10. Fill in the blanks. With 95% confidence, I estimate that the increase in log odds of smoking 100+ cigarettes for each increase by one in BMI, while holding Age constant, is between ____ and ____, on average.

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-1.2435	0.149	-8.366	0.000	-1.535	-0.952
BMI	0.0030	0.004	0.718	0.472	-0.005	0.011
Age	0.0169	0.002	10.349	0.000	0.014	0.020

○ -1.2435 and 0.149

○ 0.014 and 0.020

○ -1.535 and -0.952

◉ -0.005 and 0.011

○ Can't tell

✓ Correct