Question 1:

2*(1-pnorm(17,10,sqrt(3))) = 5.3e-05

Question 2:

2a. H_o: μ = 120
H_1: μ < 120

2b. Z-test

$$Z = \frac{\bar{X} - \mu_o}{\frac{\sigma}{\sqrt{n}}} = \frac{115 - 120}{\frac{15}{\sqrt{20}}} = -1.49$$

$$P(Z < -2.23) = 0.068$$

2c. “Fail to reject” null hypothesis; should not put drug on market!

Answer 3:

4a. X and Y might not be correlated, since the p-value isn’t as significant as we’d like. Even if X and Y are uncorrelated, this does not necessarily mean there is no relationship between X and Y.

4b. I think any of the following answers would receive full credit:

- If the effect of X on Y is very small, n=9 data points may not be enough data to detect the effect. I.e., the measurement of Y may be too noisy.
- The linear model Y = a + bX may be inappropriate. For example, the data may be quadratically related (this is what I used to generate the question) or sinusoidal, in which case the same data may produce a significantly non-0 b.
- The p-value of 0.152 is not extremely far from typical significance thresholds. It is possible that a moderate increase in the number of measurements would reveal a significant relationship.

Answer 4:

False
True
True
True
False

Answer 5:

P(state A | data) = choose(10,8)*(0.8)^8*(0.2)^2*0.5/P(data)
P(state B | data) = choose(10,8)*(0.2)^8*(0.8)^2*0.5/P(data)

P(state A | data) = 0.99976

P(state B | data) = 0.00024

n=10 or 100
pv = replicate(10000,t.test(rnorm(n,mean=1),mu=1)$p.v)
sum(pv<0.05)
pv = replicate(10000,t.test(rgamma(n,shape=1,scale=1),mu=1)$p.v)
sum(pv<0.05)

When the sample size is small, you reject a lot more for the Gamma distribution case (~10% for n=10); this problem goes away for a larger sample size, due to CLT (the normality assumption is on the mean of X).