Questions Q&A 1 – Monday Nov 1st

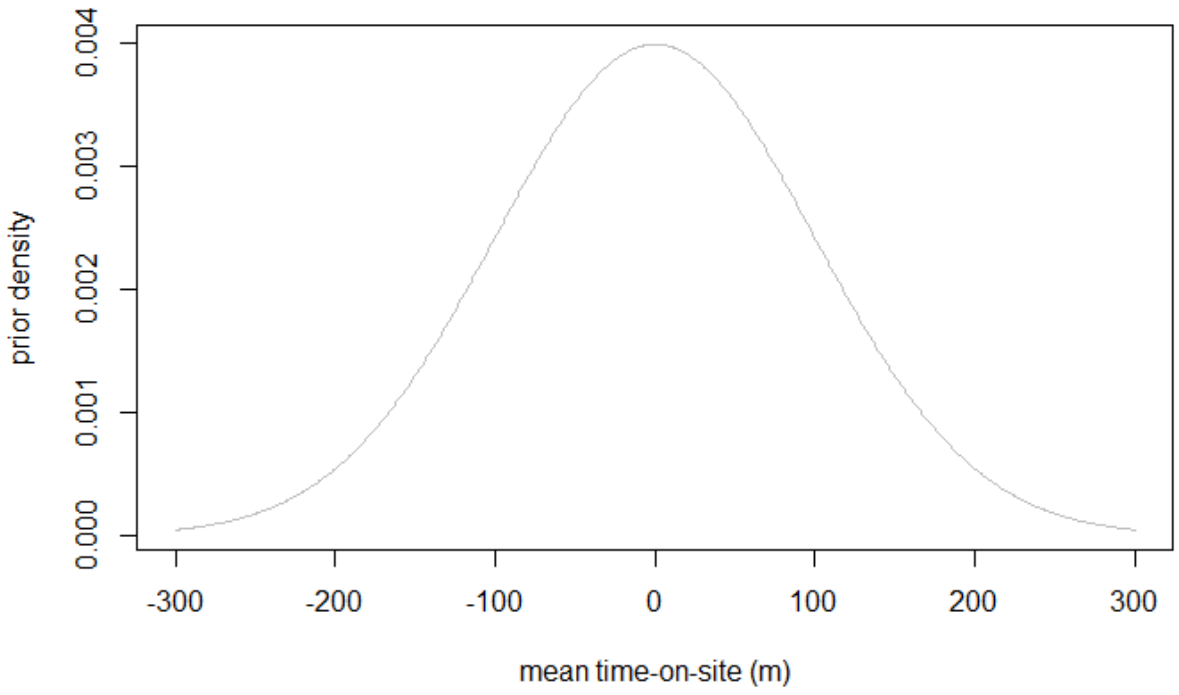
Comprehension Check Question 3 (Lab Session) – Was not covered in the lab session.

- How could you change the distribution so that you were pretty certain that the average time on site is 2 and 8 minutes?

```
# Plot the prior

plot(x=-300:300, y=dnorm(-300:300, mean=0, sd=100),

    type="l", col="gray", xlab="mean time-on-site (m)", ylab="prior density")
```



Prior and Posterior

1. How should is interpret the parameters a and b? What do they represent in the prior distribution (and thus also the posterior distribution)?

2. What is the difference between the posterior mean conversion rate and the estimated conversion rate?

3. In web clip 4 for the lab session, we compute the posterior standard deviation from the formula post_sd_B <- (1/100^2 + n_B/s^2)^-(1/2). I was wondering whether we do ^-(1/2) instead of the ^-1 that has been used in the formula below;

$$\text{mean time-on-site (m)} \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

where

$$\sigma = \left(\frac{1}{\sigma_0^2} + \frac{n}{s^2}\right)^{-1}$$

and

$$\mu = \sigma^2\left(\frac{\mu_0}{\sigma_0^2} + \frac{n\bar{y}}{s^2}\right)$$

4. In the Lab of last week, we compared two normal distributions and calculated the probability that the mean time-on-site of B is greater than the mean of A. We did this by subtracting the distributions and defined a new normal distribution using the newly found parameters. How would this work when we want to compare two beta distributions? Can we then just add and subtract the values for a and b? Or should be calculate the mean and variance of both distributions, subtract these (in the same way as for the normal) and then transform this to the values for a and b?

Bootstrap

1. When running the bootstrap, I have troubles with the "for" function; this is what I typed:

```
mub <- c()
for(b in 1:B){samp_b = sample.int(nrow(ebeer), replace=TRUE,)}
mub <- c(mub, mean(ebeer$M[samp_b]))
head(mub)
```

2. But this last head(mub) returns me the following:

   [1] 91.36043

   although in the video a lot of values appear ([1] 91.34017 90.78494 91.23552 91.21582 92.91705 92.21848)

3. With regards to the practice quiz: How are the different bootstrap equations different? When I used all of them at first, they did not give me a different answer.

Testing

1. I have a question regarding Lecture 1 Test & Roll.pptx (slide 21). Which is the formula to derive the outcome P(p_A<p_B)=0.995 ?

2. Could you explain slide 10 why the profit of 'no test' equals zero?

3. I wonder what the correct formulae is for the value of the test because on the first row it is mentioned that you have to subtract (-) the E[profit|no test] but in the row below (example) a plus sign (+) is displayed. What is the correct formulae?

Response Rate

1. How do you estimate the response rate in R (besides using the sum(mub<brk)/B, as I did not know how to derive a different answer which was asked for in the practice quiz.

2. On slide 7 of last week's lecture it says that 175 people out of 5000 people respond, and thus p_hat = 175/5000.

   What if the number of people that respond is not given and that we only know the response rates of success/failure in the past? (Just like in question 4 of the practice quiz). Is p_hat then just the expectation of these rates? So:

   p_hat = (probability of success in the past * response rate of success in the past) + (probability of failure in the past * response rate of failure in the past)

Articles

1. Do the articles contain exam material or is it an extension of the lectures and lab session? If it is exam material, do we for example have to know the formulas given in the Feit and Berman article?

2. Regarding the article for this week; the material discussed in the article is more mathematical than the information in the webclips. Therefore I was wondering whether I should memorize the formulas in the article by heart or whether I should focus more on the theoretical understanding of the paper.