2nd problem set,

1. Let P be the proportion of voters who will vote "YES". Suppose $P \sim \text{Beta}(2, 1)$. This is the prior distribution of P; i.e., the conditional distribution given all prior information. We regard P as uncertain but not random, so the probabilities are degrees-of-belief in uncertain propositions rather than relative frequencies of random events, i.e., we rely on a Bayesian interpretation of probability.

   (a) Find a 90% probability interval (a, b) for P, such that $\mathbf{Pr}(P < a) = 0.05$, $\mathbf{Pr}(P > b) = 0.05$, and so $\mathbf{Pr}(a < P < b) = 0.9$. Do not use the Bayesian Central Limit Theorem.

   15 voters are chosen randomly. Let Y be the number of those 15 who will vote "YES".

   (b) Find $\mathbf{Pr}(Y = 9 \mid P = 0.7)$. Find $\mathbf{Pr}(Y = 9)$. Find $\mathbf{E}(Y)$.

   (c) Given the observation that $Y = 9$, find the likelihood function $L(p \mid Y = 9)$.

   (d) Find the posterior probability distribution of P given that $Y = 9$, i.e., fill in the blank below:
   $$P \mid [Y = 9] \sim \boxed{\quad ? \quad}.$$

   (e) Use the Bayesian Central Limit Theorem to approximate a 90% posterior probability interval (a, b) such that $\mathbf{Pr}(P < a \mid Y = 9) = 0.05$, $\mathbf{Pr}(P > b \mid Y = 9) = 0.05$, and so $\mathbf{Pr}(a < P < b) = 0.9$.

   (f) Find the Bayes estimate $\mathbf{E}(P \mid Y = 9)$.

   Keep choosing voters randomly until you have found 9 who will vote "YES"; choose as many as it takes. Let W be the number of voters so chosen.

   (g) Given the observation that $W = 15$, find the likelihood function $L(p \mid W = 15)$.

   (h) Find the posterior probability distribution of P given that $W = 15$, i.e., fill in the blank below:
   $$P \mid [W = 15] \sim \boxed{\quad ? \quad}.$$

   (i) Discuss the relationship between the answers to (c) and (g), and the relationship between the answers to (d) and (h).

   (j) Find the Bayes estimate $\mathbf{E}(P \mid W = 15)$.

   (k) Switch from a Bayesian to a frequentist perspective. Find the maximum likelihood estimates of P given the observations in parts (c) and (g). Discuss the relationship between the two.

2. Suppose you are uncertain of a man's height H (in inches). You expressed your state of uncertainty by saying $H \sim N(71, 2.5^2)$ because you know that 71 and 2.5 are respectively the mean and the standard deviation of a population that includes this man. A very crude measuring device adds to his height a random error $\varepsilon \sim N(0, 1)$, so that the height measured with error is $M = H + \varepsilon$, and H and ε are independent.

CONTINUED⟶

(a) Write the conditional distribution of M given H in the form

$$\text{“}M \mid [H = h] \sim N(?,\ ?)\text{”}$$

— fill in the blanks. (Don't use Bayes' formula! That is wrong. This is an easier question than that.) Write the likelihood function

$$L(h) = f_{M|H=h}(m)$$

in the form

$$[\text{constant}] \cdot e^{(-1/2)(\ ?\ )^2}$$

— fill in the blank and don't worry at this point about the value of the "constant."

(b) Multiply the prior probability density function by the likelihood function and get the posterior probability density function in the form

$$f_{H|M=m}(h) = [\text{constant}] \cdot e^{(-1/2)(\ ?\ )^2}$$

— fill in the blank with something that looks like this:

$$\frac{h - \text{something}}{\text{something}}$$

where the two "somethings" do not depend on h. (You may need to do some algebraic massaging of the exponent to make it look like that.)

(c) Use the answer to (b) to find the Bayes estimator

$$\mathbf{E}(H \mid M = m) \quad = \quad \begin{array}{l}\text{a weighted average of the observed measurement } m \text{ (i.e., the} \\ \text{height-measured-with-error) and the prior expected value.}\end{array}$$

Specify the weights in the weighted average. "Use the answer to (b)" means do it by that method and not by some other method. This can be done very quickly since you've already done part (b).

(d) Suppose you observe $M = 74$. Find a 90% posterior probability interval (h_1, h_2), so that $\Pr(H < h_1 \mid M = 74) = 0.05$, $\Pr(H > h_2 \mid M = 74) = 0.05$, and so $\Pr(h_1 < H < h_2 \mid M = 74) = 0.9$.