Answers to the 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.

   Answer: When I wrote "Do the problem numerically rather than algebraically" I actually had in mind beta distributions with much larger values of α and β; when $\alpha = 2$ and $\beta = 1$ the problem is easily tractable algebraically. We have:

$$\begin{aligned} f_P(p) &= [\text{constant}] \cdot p^{2-1}(1-p)^{1-1} = 2p. \\ 0.05 &= \mathbf{Pr}(P < a) = \int_0^a 2p\,dp = a^2, \text{ and} \\ 0.95 &= \mathbf{Pr}(P < b) = \int_0^b 2p\,dp = b^2, \end{aligned}$$

   so $a = \sqrt{0.05} \cong 0.2236068$ and $b = \sqrt{0.95} = 0.974679$.

   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 $\mathbf{E}(Y)$.

   Answer: Since $Y \mid [P = 0.7] \sim \text{Bin}(15, 0.7)$ we have

$$\mathbf{Pr}(Y = 9 \mid P = 0.7) = \binom{15}{9} 0.7^9 0.3^6 = 0.1472\ldots.$$

   Next we have

$$\begin{aligned} \mathbf{Pr}(Y = 9) &\overset{\boxed{\text{Law of Total Probability}}}{=} \mathbf{E}(\mathbf{Pr}(Y = 9 \mid P)) = \mathbf{E}\left(\binom{15}{9} P^9 (1-P)^6\right) \\ &= \int_0^1 \binom{15}{9} p^9 (1-p)^6 f_P(p)\,dp = \int_0^1 \binom{15}{9} p^9 (1-p)^6\, 2p\,dp \\ &= 2\binom{15}{9} \int_0^1 p^{11-1}(1-p)^{7-1}\,dp = 2 \cdot \frac{15!}{9!6!} \cdot \frac{\Gamma(11)\Gamma(7)}{\Gamma(18)} \\ &= 2 \cdot \frac{15!}{9!6!} \cdot \frac{10!6!}{17!} = 2 \cdot \frac{10}{16 \cdot 17} = \frac{5}{4 \cdot 17} \cong 0.0735294\ldots. \end{aligned}$$

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

<u>**Answer**</u>**:**

$$L(p \mid Y = 9) = \mathbf{Pr}(Y = 9 \mid P = p) = \binom{15}{9} p^9 (1-p)^6. \textbf{ (That's all!)}$$

Several of you made this more complicated than it is, by thinking that what was asked for was the posterior probability density. That is an error.

(d) Find the posterior probability distribution of P given that $Y = 9$, i.e., fill in the blank below:

$$P \mid [Y = 9] \sim \boxed{\quad ? \quad}.$$

<u>**Answer**</u>**:**

$$\begin{aligned} f_{P|Y=9}(p) &\overset{\boxed{\text{Bayes' formula}}}{\downarrow}{=} [\text{constant}] \cdot f_P(p) \cdot L(p \mid Y = 9) \\ &= [\text{constant}] \cdot 2p \cdot \binom{15}{9} p^9 (1-p)^6 \\ &= [\text{constant}] \cdot p^{11-1}(1-p)^{7-1}. \end{aligned}$$

$$\text{Therefore } P \mid [Y = 9] \sim \text{Beta}(11, 7).$$

(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 $\mathbf{Pr}(a < P < b) = 0.9$.

<u>**Answer**</u>**:** You were told to use both of the versions of the Bayesian CLT that were stated in the handout of February 19th. In both versions you need the fact that $\Phi(-1.645) = 0.05$ and $\Phi(1.645) = 0.95$, where Φ is the standard normal c.d.f.

First version: Since the posterior distribution is $\text{Beta}(11, 7)$, the posterior expectation is $\dfrac{11}{11+7} = 0.61111\ldots$, and the posterior standard deviation is

$$\sqrt{\frac{11 \cdot 7}{(11+7)^2(11+7+1)}} = 0.111839716\ldots$$

so the endpoints of the desired interval are

$$\frac{11}{11+7} \pm 1.645 \sqrt{\frac{11 \cdot 7}{(11+7)^2(11+7+1)}} \cong 0.6111111 \pm 0.1839763334\ldots$$

and so the interval is $(0.42713478\ldots, 0.79508744\ldots)$.

Second version: The posterior density is

$$f(p) = [\text{constant}] \cdot p^{10}(1-p)^6 \text{ for } 0 < p < 1$$

and $= 0$ if $p < 0$ or $p > 1$, and the posterior mode is the value of p that maximizes that function. We have

$$\frac{d}{dp}\log\left(p^{10}(1-p)^6\right) = \frac{10}{p} - \frac{6}{1-p} = \frac{-16(p-5/8)}{p(1-p)} \begin{cases} > 0 & \text{if } 0 < p < 5/8, \\ = 0 & \text{if } p = 5/8, \\ < 0 & \text{if } 5/8 < p < 1. \end{cases}$$

Consequently the posterior mode is at $p = 5/8 = 0.625$ (this is not a maximum likelihood estimate, since what was being maximized was not the likelihood function). Next we need $d = (-k''(5/8))^{-1/2}$ where $k(p) = \log f(p)$. So we have

$$\begin{aligned}
k(p) &= \log f(p) = [\text{constant}] + 6\log(p) + 10\log(1-p) \\
k'(p) &= \frac{6}{p} - \frac{10}{1-p} \\
k''(p) &= -\frac{6}{p^2} - \frac{10}{(1-p)^2} \\
k''(5/8) &= -\frac{384}{25} - \frac{640}{9} = -\frac{384\cdot 9 + 640\cdot 25}{25\cdot 9} = -\frac{19456}{225} = -86.471111\ldots \\
\sqrt{-k''(5/8)} &= 0.107538625\ldots.
\end{aligned}$$

So the endpoints of the desired interval are

$$0.625 \pm 1.645\,(0.107538625\ldots) = 0.44809896\ldots \text{ and } 0.801901\ldots$$

(So maybe you would want a larger sample than this before you regard the Bayesian Central Limit Theorem as giving sufficiently accurate results.)

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

<u>Answer</u>:

Since $P \mid [Y = 9] \sim \text{Beta}(11, 7)$ we must have $\mathbf{E}(P \mid Y = 9) = \dfrac{11}{11+7} = 0.61111\ldots.$

CONTINUED⟶

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 $L(p \mid W = 15)$.

Answer: The number W of independent trials until a specified number of successes (in this case 9) with probability P of success on each trial, has a **negative binomial distribution**, and we write $W \sim \text{Negbin}(9, P)$. Here's a very terse review:

$$\begin{aligned} & \mathbf{Pr}(W = w) \\ = \; & \mathbf{Pr}(9-1 \text{ successes on the first } w-1 \text{ trials and success on the } w\text{th trial}) \\ = \; & \binom{w-1}{9-1} p^{9-1}(1-p)^{w-9} \cdot p \\ = \; & \binom{w-1}{9-1} p^{9}(1-p)^{w-9} \end{aligned}$$

for $w \in \{9, 10, 11, 12, \dots\}$ (obviously $W \geq 9$ with probablity 1).
Back to the problem at hand:

$$L(p \mid W = 15) \;=\; \mathbf{Pr}(W = 15 \mid P = p) = \binom{15-1}{9-1} p^9(1-p)^6.$$

(h) Find the posterior probability distribution of P given that $W = 15$, i.e., fill in the blank below:

$$P \mid [W = 15] \sim \boxed{\quad ? \quad}.$$

Answer: Procede as in part (d):

$$\begin{aligned} f_{P|Y=9}(p) \;&\overset{\boxed{\text{Bayes' formula}}}{=}\; [\text{constant}] \cdot f_P(p) \cdot L(p \mid Y = 9) \\ &=\; [\text{constant}] \cdot 2p \cdot \binom{14}{8} p^9(1-p)^6 \\ &=\; [\text{constant}] \cdot p^{11-1}(1-p)^{7-1}. \end{aligned}$$

$$\text{Therefore } P \mid [Y = 9] \;\sim\; \text{Beta}(11, 7).$$

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

Answer: Several students answered that the answers in (c) and (g) are the same. They are not. We have:

$$\begin{aligned} \text{in part (c):} \quad L(p \mid Y = 9) &= \mathbf{Pr}(Y = 9 \mid P = p) = \binom{15}{9} p^9 (1-p)^6, \\ \text{and in part (d):} \quad L(p \mid W = 15) &= \mathbf{Pr}(W = 15 \mid P = p) = \binom{15-1}{9-1} p^9 (1-p)^6. \end{aligned}$$

Since $\binom{15}{9}/\binom{14}{8} = 3/5 \neq 1$ the two likelihood functions are not equal. But each is a constant multiple of the other, or in other words, they are proportional.

The answers to (d) and (h) are identical. Proportional likelihood functions that are not equal are equivalent in the application of Bayes' formula.

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

Answer: Since $P \mid [W = 15] \sim \text{Beta}(11, 7)$, it follows that $\mathbf{E}(P \mid W = 15) = 11/(11+7) = 0.61111\ldots.$

(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.

Answer: In each case we have

$$L(p) = [\text{constant}] \cdot p^9 (1-p)^6.$$

The constant factor in part (c) differs from that in part (g). Since the constant factor is positive in both cases the value of p that maximizes $L(p)$ is the same in both cases. Letting $\ell(p) = \log L(p)$, we have

$$\ell'(p) = \frac{d}{dp}(6 \log p + 9 \log(1-p)) = \frac{6}{p} - \frac{9}{1-p} = \frac{-15(p - 3/5)}{p(1-p)} \begin{cases} > 0 & \text{if } 0 < p < 3/5, \\ = 0 & \text{if } p = 3/5, \\ < 0 & \text{if } 3/5 < p < 1. \end{cases}$$

So in both cases, the MLE is $\widehat{p} = 3/5$.

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."

Answer: Since H and ε are independent, the distribution of ε given $H = h$ is no different from the marginal distribution of ε; it is $N(0, 1)$. Given the condition $H = h$, the random variable H turns into a constant, so we have

$$\begin{aligned} \varepsilon + H \mid [H = h] &\sim N(h, 1^2), \\ \text{i.e.,} \qquad M \mid [H = h] &\sim N(h, 1^2). \end{aligned}$$

The likelihood is

$$L(h) = f_{M|H=h}(m) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2}\left(\frac{m-h}{1}\right)^2\right) = [\text{constant}] \cdot \exp\left(\frac{-1}{2}(m-h)^2\right).$$

(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.)

Answer: The prior density is

$$[\text{constant}] \cdot \exp\left(\frac{-1}{2}\left(\frac{h-71}{2.5}\right)^2\right).$$

The likelihood is as reported in part (a) above. Multiplying these, we get

$$[\text{constant}] \cdot \exp\left(\frac{-1}{2}\left((m-h)^2 + \left(\frac{h-71}{2.5}\right)^2\right)\right).$$

Remembering that "constant" in this context means "not depending on m," we simplify the quadratic polynomial in h:

$$
\begin{aligned}
&(m-h)^2+\left(\frac{h-71}{2.5}\right)^2 \\[2ex]
&= \left(\frac{2.5^2+1}{2.5^2}\right)h^2-2\left(m+\frac{71}{2.5^2}\right)h+\begin{pmatrix}\text{terms not}\\ \text{depending}\\ \text{on } h\end{pmatrix} \\[2ex]
&= \left(\frac{2.5^2+1}{2.5^2}\right)\left(h^2-2\left(\frac{2.5^2\cdot m+1\cdot 71}{2.5^2+1}\right)h\right)+\begin{pmatrix}\text{terms not}\\ \text{depending}\\ \text{on } h\end{pmatrix} \\[2ex]
&= \left(\frac{2.5^2+1}{2.5^2}\right)\left(h^2-2\left(\frac{2.5^2\cdot m+1\cdot 71}{2.5^2+1}\right)h+\left(\frac{2.5^2\cdot m+1\cdot 71}{2.5^2+1}\right)^2\right)+\begin{pmatrix}\text{terms not}\\ \text{depending}\\ \text{on } h\end{pmatrix} \\[2ex]
&= \left(\frac{2.5^2+1}{2.5^2}\right)\left(h-\left(\frac{2.5^2\cdot m+1\cdot 71}{2.5^2+1}\right)\right)^2+\begin{pmatrix}\text{terms not}\\ \text{depending}\\ \text{on } h\end{pmatrix} \\[2ex]
&= \left(\frac{h-\left(\frac{2.5^2\cdot m+1\cdot 71}{2.5^2+1}\right)}{\left(\frac{2.5}{\sqrt{2.5^2+1}}\right)}\right)^2+\begin{pmatrix}\text{terms not}\\ \text{depending}\\ \text{on } h\end{pmatrix}.
\end{aligned}
$$

Consequently we have

$$
H \mid [M=m] \sim N\left(\frac{2.5^2\cdot m+1\cdot 71}{2.5^2+1},\quad \frac{2.5^2}{2.5^2+1}\right).
$$

The posterior expected value is a weighted average of the measured value m and the prior expected value 71, with weights inversely proportional to 1 and 2.5^2 — the two variances.

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

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

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).

<u>**Answer**</u>**:**

$$\begin{aligned} \mathbf{E}(H \mid M = m) &= \frac{2.5^2 \cdot m + 1 \cdot 71}{2.5^2 + 1} = \left(\frac{2.5^2}{2.5^2 + 1^2}\right) m + \left(\frac{1^2}{2.5^2 + 1^2}\right) 71 \\ &= w_1 \cdot m + w_2 \cdot 71 \end{aligned}$$

where the weights w_1 and w_2 are the numbers in "(" parentheses ")" on the previous line. In order that this be a weighted average it is necessary and sufficient that the weights be nonnegative and that their sum be 1.

Observe that

$$\begin{aligned} w_1 \cdot m + w_2 \cdot 71 &= \left(\frac{2.5^2}{2.5^2 + 1^2}\right) m + \left(\frac{1^2}{2.5^2 + 1^2}\right) 71 \\ &= k \left[\left(\frac{1}{1^2}\right) m + \left(\frac{1}{2.5^2}\right) 71\right] \\ &= k \left[\frac{1}{\mathbf{var}(\varepsilon)} m + \frac{1}{\mathbf{var}(H)} 71\right] \end{aligned}$$

(you can quickly find the value of k). Thus, the weights assigned to (1) the measurement, and (2) the prior average, are inversely proportional to (1) the variance of the measurement error, and (2) the prior variance, respectively.

CONTINUED⟶

(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 $\Pr(h_1 < H < h_2 \mid M = 74) = 0.9$.

Answer: From part (b) we conclude that

$$\begin{aligned}
\mathbf{E}(H \mid M = 74) &= \left(\frac{2.5^2}{2.5^2 + 1^2}\right) 74 + \left(\frac{1^2}{2.5^2 + 1^2}\right) 71 = 73.5862\ldots. \\
\mathbf{var}(H \mid M = 74) &= \frac{2.5^2}{2.5^2 + 1^2} = 0.862\ldots.
\end{aligned}$$

Since we know that $\Phi(-1.645) = 0.05 = 1 - \Phi(+1.645)$, we need the interval whose endpoints are

$$\begin{aligned}
\text{expectation} \pm 1.645\,\mathbf{SD} &= 73.5862\ldots \pm 1.645(2.5)/\sqrt{2.5^2 + 1} \\
&= 73.5862\ldots \pm 1.527344 \\
&= 72.05886, \quad 75.11355.
\end{aligned}$$

Therefore we have

$$\begin{aligned}
\mathbf{Pr}(72.05886 < H < 75.11355 \mid M = 74) &= 0.95. \\
\mathbf{Pr}(H < 72.05886 \mid M = 74) &= 0.05. \\
\mathbf{Pr}(H > 75.11355 \mid M = 74) &= 0.05.
\end{aligned}$$