▷ Exercise 2.8.[2] Assuming a uniform prior on f_H, $P(f_H) = 1$, solve the problem posed in example 2.7 (p.30). Sketch the posterior distribution of f_H and compute the probability that the $N+1$th outcome will be a head, for

(a) $N = 3$ and $n_H = 0$;
(b) $N = 3$ and $n_H = 2$;
(c) $N = 10$ and $n_H = 3$;
(d) $N = 300$ and $n_H = 29$.

Solution 2.8:

Posterior = $\frac{likelihood\ x\ prior}{evidence}$

Prior: $P(f_H) = 1$

Evidence: $P(n_H \mid N) = \int_0^1 f_h{}^{n_H}(1-f_h)^{N-n_H})\, df_h$

What we obtain from the Beta Integral according to n_H and N is:

$P(n_H \mid N) = \frac{n_H!(N-n_H)!}{(n_H+(N-n_H)+1)!} = \frac{n_H!(N-n_H)!}{(N+1)!}$

Likelihood: $P(n_H \mid f_H, N) = f_h{}^{n_H}(1-f_h)^{N-n_H})$

According to this we can calculate posterior probability as shown below:

$P(f_H \mid n_H, N) = \frac{f_h{}^{n_H}(1-f_h)^{N-n_H})\ x\ 1}{\frac{n_H!(N-n_H)!}{(N+1)!}}$

Hypothesis says that toss N+1 will result *head,* we will integrate over f_h by using sum rule:

$P(h \mid n_H, N) = \int P(h \mid f_H) . P(f_H \mid n_H, N) df_h$

$= \int_0^1 P(h \mid f_H). \frac{f_h{}^{n_H}(1-f_h)^{N-n_H})}{\frac{n_H!(N-n_H)!}{(N+1)!}} df_h$

Since $P(h \mid f_H) = f_H$;

$P(h \mid n_H, N) = \int_0^1 \frac{f_h{}^{n_H}(1-f_h)^{N-n_H})}{\frac{n_H!(N-n_H)!}{(N+1)!}} df_h$

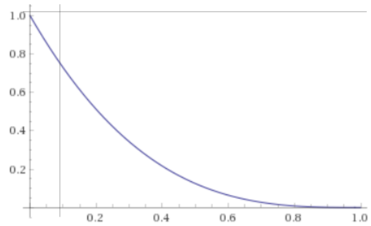
$= \frac{(N+1)!}{n_H!(N-n_H)!} \int_0^1 f_h{}^{n_H}(1-f_h)^{N-n_H})\, df_h$

When we use Beta Integral one more time:

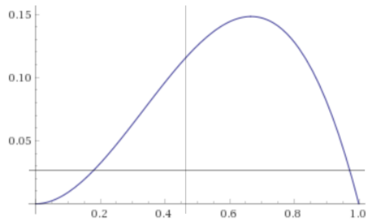
$P(h \mid n_H, N) = \frac{(N+1)!}{n_H!(N-n_H)!} \cdot \frac{(n_H+1)!(N-n_H)!}{(N+2)!}$

$= \frac{n_H+1}{N+2}$

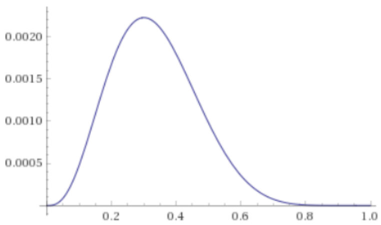
a) N = 3 and $n_H = 0$; $\frac{0+1}{3+2} = {}^{1}/_{5}$



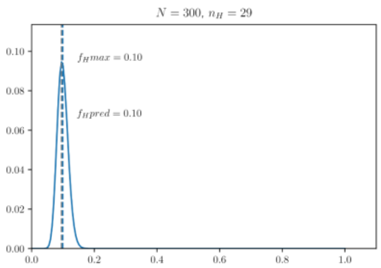
b) N = 3 and $n_H = 2$; $\frac{2+1}{3+2} = {}^{3}/_{5}$



c) N = 10 and $n_H = 3$; $\frac{3+1}{10+2} = {}^{1}/_{3}$



d) N = 300 and $n_H = 29$; $\frac{29+1}{300+2} = {}^{30}/_{302}$



Please solve the following two exercises.

Example 2.10. Urn A contains three balls: one black, and two white; urn B contains three balls: two black, and one white. One of the urns is selected at random and one ball is drawn. The ball is black. What is the probability that the selected urn is urn A?

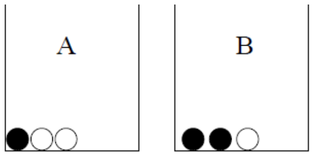


Figure 2.7. Urns for example 2.10.

Solution 2.10:

$\mathrm{P}(Urn\ A \mid black)$ = ?

Bayes Equation for the given problem:

$$\mathbf{P}(\boldsymbol{Urn\ A} \mid \boldsymbol{black}) = \frac{\boldsymbol{P}(\boldsymbol{black} \mid \boldsymbol{Urn\ A})\ \boldsymbol{x}\ \boldsymbol{P}(\boldsymbol{Urn\ A})}{\boldsymbol{P}(\boldsymbol{black})}$$

Probability for selecting one of the Urns (A or B):

$$\mathrm{P}(Urn\ A) = \mathrm{P}(Urn\ B) = 0.5$$

Probability for drawn black ball from each Urn separately:

$$\mathrm{P}(black) = \left({}^{1}\!/_{3}\ x\ 0.5 \right) + \left({}^{2}\!/_{3}\ x\ 0.5 \right) = {}^{1}\!/_{2}$$

Solving the Bayes equation:

$$\frac{P(black \mid Urn\ A)\ x\ P(Urn\ A)}{P(black)} = \frac{{}^{1}\!/_{3}\ x\ {}^{1}\!/_{2}}{{}^{1}\!/_{2}} = \mathbf{{}^{1}\!/_{3}}$$

Example 2.11. Urn A contains five balls: one black, two white, one green and one pink; urn B contains five hundred balls: two hundred black, one hundred white, 50 yellow, 40 cyan, 30 sienna, 25 green, 25 silver, 20 gold, and 10 purple. [One fifth of A's balls are black; two-fifths of B's are black.] One of the urns is selected at random and one ball is drawn. The ball is black. What is the probability that the urn is urn A?

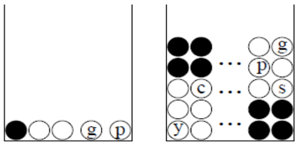


Figure 2.8. Urns for example 2.11.

Solution 2.11:

$\mathrm{P}(Urn\ A \mid black) = ?$

Bayes Equation for the given problem:

$$\boldsymbol{P(Urn\ A \mid black) = \frac{P(black \mid Urn\ A)\ x\ P(Urn\ A)}{P(black)}}$$

Probability for selecting one of the Urns (A or B):

$$\mathrm{P}(Urn\ A) = \mathrm{P}(Urn\ B) = 0.5$$

Probability for black ball drawn from Urn A:

$$P(black \mid Urn\ A) = {}^{1}/_{5}$$

Probability for selecting black from any of the Urns:

$$\mathrm{P}(black) = P(black \mid Urn\ A) + P(black \mid Urn\ B)$$

$$\mathrm{P}(black) = \left({}^{1}/_{5}\ x\ 0.5\right) + \left({}^{2}/_{5}\ x\ 0.5\right) = {}^{3}/_{10}$$

Solving the Bayes equation:

$$\mathrm{P}(Urn\ A \mid black) = \frac{P(black \mid Urn\ A)\ x\ P(Urn\ A)}{P(black)} = \frac{{}^{1}/_{5}\ x\ {}^{1}/_{2}}{{}^{3}/_{10}} = \mathbf{{}^{1}/_{3}}$$

Exercise 3.5.[2, p.59] Sketch the posterior probability $P(p_a \mid s = aba, F = 3)$. What is the most probable value of p_a (i.e., the value that maximizes the posterior probability density)? What is the mean value of p_a under this distribution?

Answer the same questions for the posterior probability $P(p_a \mid s = bbb, F = 3)$.

Solution 3.5:

Posterior probability:

$P(P_a \mid s = 'aba', F = 3)$

If we extract *s* over posterior probability:

${P_a}^{F_a}(1 - P_a)^{F_b})$

'aba' → $P_a.(1 - P_a).P_a$

$$P(P_a \mid s, F) = \frac{{P_a}^{F_a}(1 - P_a)^{F_b}}{\frac{F_b! \text{ x } F_b!}{(F_b! + F_b! + 1)!}}$$

If we add parameters to equation above considering that $F_a = 2$ and $F_b = 1$:

$$\frac{{P_a}^2(1 - P_a)^1}{\frac{2! \text{ x } 1!}{(2 + 1 + 1)!}}$$

Here we must set derivative of the equation to 0 in order to find most expectation (uniform distribution has an expectation but no most probable value)[2]:

a)

${}^{4!}/_{2!}$ x ${P_a}^2(1 - P_a)^1 \, d/dP_a$ = 0

12 x $({P_a}^2 - {P_a}^3)d/dP_a$ = 0

12 x $(2P_a - 3P_a)$ = 0

According to this the value that maximizes posterior probability:

$P_a = {}^{2}/_{3}$

In order to calculate mean, we must use equation in (3.13)[2]:

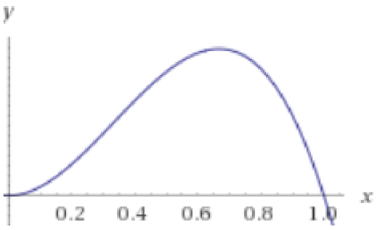
$\mu = 12$ x $\int_0^1 P_a.{P_a}^2(1 - P_a)^1 \, dP_a$ = 12 x $\int_0^1 {P_a}^3 - {P_a}^4 \, dP_a$

If we solve integral:

μ = 12 x $(\frac{{P_a}^4}{4} - \frac{{P_a}^5}{5} \; for \; [0,1] = 12$ x $\left(\frac{1}{4} - \frac{1}{5}\right)$

μ = $^{3}/_{5}$

Sketch of posterior probability:



b)

Posterior probability:

$P(P_a \mid s = 'bbb, F = 3)$

If we extract *s* over posterior probability:

$P_a{}^{F_a}(1 - P_a)^{F_b})$

'bbb' → $(1 - P_a).(1 - P_a).(1 - P_a)$

$$P(P_a \mid s, F) = \frac{P_a{}^{F_a}(1 - P_a)^{F_b}}{\frac{F_b! \text{ x } F_b!}{(F_b! + F_b! + 1)!}}$$

If we add parameters to equation above considering that $F_a = 0$ and $F_b = 3$:

$$\frac{P_a{}^0(1 - P_a)^3}{\frac{0! \text{ x } 3!}{(0 + 3 + 1)!}}$$

Here we must set derivative of the equation to 0 in order to find expectation (uniform distribution has an expectation but no most probable value)[2]:

$^{4!}/_{3!}$ x $P_a{}^0(1 - P_a)^3 \ d/dP_a$

$4 \text{ x } (1 - P_a)^3 d/dP_a$

$4 \text{ x } (-3(P_a - 1)^2)$

Here as long as we do not know if the distribution is convex or concave, we cannot use the equation below. The value of x that maximizes the expectation is 0. We can also confirm this by setting P_a = 0:

$P_a = 0$

$P_a{}^0(1 - P_a)^3$ = $0^0(1 - 0)^3$ = $\underline{1}$

1 is the maximum value for any probability. So, this confirms that our answer is correct.

In order to calculate mean, we must use equation in (3.13)[2]:

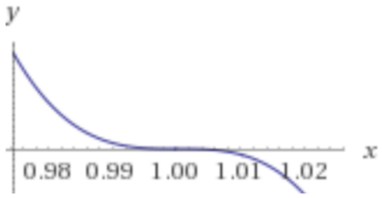
$\mu = 4 \text{ x} \int_0^1 P_a(1 - P_a)^3 \ dP_a$

If we solve integral:

μ = 4 x $\left(\frac{P_a{}^2}{2} - \frac{3P_a{}^3}{3} + \frac{3P_a{}^4}{4} - \frac{P_a{}^5}{5}\right.$ $for\ [0,1] = 4$ x $\left(\frac{1}{2} - 1 + \frac{3}{4} - \frac{1}{5}\right)$

μ = $^1/_5$

Sketch of posterior probability:



▷ Exercise 3.12.[2] A bag contains one counter, known to be either white or black. A white counter is put in, the bag is shaken, and a counter is drawn out, which proves to be white. What is now the chance of drawing a white counter? [Notice that the state of the bag, after the operations, is exactly identical to its state before.]

Solution 3.12:

Here we have two hypotheses: let *H* = 0 mean that he original counter in the bag was white and *H* = 1 that it was black. In our case prior probabilities are equal. The bag contained a white counter and the unknown one. The data tells us that when a randomly selected counter was drawn from the bag, it turned out to be white. The probability of this result according to each hypothesis:

$P(D \mid H = 0) \quad \equiv \quad P(White\ Counter \mid Drawn\ White) \ = 1$

$P(D \mid H = 1) \quad \equiv \quad P(Black\ Counter \mid Drawn\ White) = \ ^1/_2$

According to these we can calculate probability for grabbing white for each hypothesis:

$P(Drawn\ White) =$ (1 x 1/2) + (1/2 x 1/2) = 3/4

Posterior probability after counter is selected according to Bayes' theorem as it follows:

$P(White\ Counter \mid Drawn\ White)$ = $\frac{P(Drawn\ White \mid White\ Counter)\ x\ P(White\ Counter)}{P(Drawn\ White)}$

= $\frac{1 \text{ x } 1/2}{3/4}$

= 2/3

$P(Black\ Counter \mid Drawn\ White)$ = $\frac{P(Drawn\ White \mid Black\ Counter)\ x\ P(Black\ Counter)}{P(Drawn\ White)}$

$$= \frac{1/2 \times 1/2}{3/4}$$

$$= 1/3$$

▷ Exercise 3.14.[1] In a game, two coins are tossed. If either of the coins comes up heads, you have won a prize. To claim the prize, you must point to one of your coins that is a head and say 'look, that coin's a head, I've won'. You watch Fred play the game. He tosses the two coins, and he points to a coin and says 'look, that coin's a head, I've won'. What is the probability that the *other* coin is a head?

Here we have four cases for H = Head, T = Tail. What we know is Fred claims that he won in situations of {HT, TH, HH}, so that we can ignore case {TT} for the situations that Fred wins.

$$P(Other\ Toss'is\ tail \mid Fred\ claims\ that\ he\ won) = \frac{P(Fred\ Wins \mid Other\ Toss{=}T)\ x\ P(Toss{=}T)}{P(Fred\ Wins)}$$

$$= \frac{{}^{2}/_{3} \times {}^{3}/_{4}}{{}^{3}/_{4}}$$

$$= {}^{2}/_{3}$$

$$P(Other\ Toss'is\ head \mid Fred\ claims\ that\ he\ won) = \frac{P(Fred\ Wins \mid Other\ Toss{=}H)\ x\ P(Toss{=}H)}{P(Fred\ Wins)}$$

$$= \frac{{}^{1}/_{3} \times {}^{3}/_{4}}{{}^{3}/_{4}}$$

$$= {}^{1}/_{3}$$

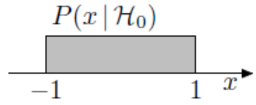
Exercise 28.1.[3] Random variables x come independently from a probability distribution $P(x)$. According to model $\mathcal{H}_0$, $P(x)$ is a uniform distribution

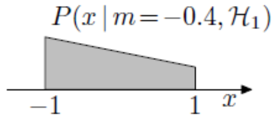
$$P(x \mid \mathcal{H}_0) = \frac{1}{2} \qquad x \in (-1, 1). \tag{28.20}$$

According to model $\mathcal{H}_1$, $P(x)$ is a nonuniform distribution with an unknown parameter $m \in (-1, 1)$:

$$P(x \mid m, \mathcal{H}_1) = \frac{1}{2}(1 + mx) \qquad x \in (-1, 1). \tag{28.21}$$

Given the data $D = \{0.3, 0.5, 0.7, 0.8, 0.9\}$, what is the evidence for $\mathcal{H}_0$ and $\mathcal{H}_1$?





Solution 28.1:

H_0)

By using given parameters;

D = { 0.3, 0.5, 0.7, 0.8, 0.9 }, m = unknown

Bayes probability of H_0 could be calculated as shown below.

$$P(H_0 \mid D) = \frac{P(D| \, H_0\,) \text{ x } P(H_0\,)}{P(D)}$$

Given prior probability for our distribution:

$P(H_0)$ = $^1/_2$

Here we have uniform distribution, which means that we have same probability for each element in our dataset. So calculating probability for 5 times would give us the correct answer:

$$P(D| \, H_0\,) = \prod_{1}^{5} \frac{1}{2}$$

$$P(D| \, H_0\,) = \frac{1^5}{2} \; x \; (1 - \frac{1}{2})^0 = \frac{1}{32} = 0.03125$$

H_1)

Bayes probability of H_1 could be calculated as shown below.

$$P(H_1 \mid D, m) = \frac{P(D| \, H_1\,, m) \text{ x } P(H_1\,)}{P(D| \, m)}$$

Considering the fact that m is uniform, prior probability for *m* should be:

$P(H_1\,)$ = $^1/_2$

Since “the evidence for a model is usually the normalizing constant of an earlier Bayesian inference”[2], and also according to equation 3.20 from the book[2] and sample equation[1] , so because of we have independent(discreate) values, we can formulate H_1 according to *D* sample of data and *m* parameter like this:

$$P(D|\,H_1\,,m) = \prod_{i=1}^{5} \frac{1}{2} \text{ x } (1 + m.x_i) \quad \forall x_i \in D$$

$$P(D|\,H_1\,,m) = \left(\tfrac{1}{2} \text{ x } (1 + 0.3m)\right) * \left(\tfrac{1}{2} \text{ x } (1 + 0.5m)\right) * \left(\tfrac{1}{2} \text{ x } (1 + 0.7m)\right) * \left(\tfrac{1}{2} \text{ x } (1 + 0.8m)\right) * \left(\tfrac{1}{2} \text{ x } (1 + 0.9m)\right)$$

We need to solve this equation for each possible value of m. So that we must take integral of the equation in [-1,1]:

$$P(D|\,H_1\,,m) = \int_{-1}^{1} \left(\frac{1}{2} \text{ x } (1 + 0.3m)\right) * \left(\frac{1}{2} \text{ x } (1 + 0.5m)\right) * \left(\frac{1}{2} \text{ x } (1 + 0.7m)\right) * \left(\frac{1}{2} \text{ x } (1 + 0.8m)\right) * \left(\frac{1}{2} \text{ x } (1 + 0.9m)\right) dm$$

$P(D|\,H_1\,,m) =$ 0.15403

Evidence for H_0 = 0.03125

Evidence for H_1 = 0.15403

Referances

1-Rabe-Hesketh, S. (2000). Maximum Likelihood Estimation. *Handbook of Statistical Analyses Using Stata, Fourth Edition*, 1–7. https://doi.org/10.1201/9781584888574.ch13

$$L(\theta) = \prod_{i=1}^{n} P(X_i|\theta)$$

2- Andrew, A. M. (2004). Information Theory, Inference, and Learning Algorithms. In *Kybernetes* (Vol. 33). https://doi.org/10.1108/03684920410534506