Bayes School 2013 at Stellenbosh

Problems based on lectures by Giulio d'Agostini

Solutions

1. NONE (see e.g. en.wikipedia 'wiki/P-value#Misunderstandings')

2. **Answer: 2/3**

   Detailed solution: Calling A the hypothesis "white ball extracted from the big box" and B "black ball extracted from the big box", with $P(A\,|\,I) = P(B\,|\,I) = 1/2$, after having introduced a white ball in the bag, the probabilities to extract the first white ball from the little bag are

   $$\begin{aligned} P(W^{(1)}\,|\,A,I) &= 1 \\ P(W^{(1)}\,|\,B,I) &= 1/2\,. \end{aligned}$$

   Using Bayes theorem the probabilities of the two hypotheses, conditioned by $W^{(1)}$ are then

   $$\begin{aligned} P(A\,|\,W^{(1)},I) &= \frac{P(W^{(1)}\,|\,A,I)\cdot P(A\,|\,I)}{P(W^{(1)}\,|\,A,I)\cdot P(A\,|\,I) + P(W^{(1)}\,|\,B,I)\cdot P(B\,|\,I)} \\ &= \frac{1\times 1/2}{1\times 1/2 + 1/2\times 1/2} = \frac{2}{3} \\ P(B\,|\,W^{(1)},I) &= \ldots = \frac{1}{3} \quad \left[= 1 - P(A\,|\,W^{(1)},I)\right]\,, \end{aligned}$$

   which are obviously equal to the probabilities to have white or black in a second extraction from the bag.

3. (a) **Answer: 11.3%**

   Detailed solution: [Assuming hereafter implicit the 'background information' I, i.e. writing $P(T)$ for $P(T\,|\,I)$, etc.]

   $$\begin{aligned} P(T) &= P(T\,|\,\mu)\,P_\circ(\mu) + P(T\,|\,\pi)\,P_\circ(\pi) \\ &= 0.95\times 0.02 + 0.02\times 0.9 = 0.113 \end{aligned}$$

   (b) **Answer: 84%**

   Detailed solution:

   $$\begin{aligned} P(\mu\,|\,T) &= \frac{P(T\,|\,\mu)\,P_\circ(\mu)}{P(T)} \\ &= \frac{P(T\,|\,\mu)\,P_\circ(\mu)}{P(T\,|\,\mu)\,P_\circ(\mu) + P(T\,|\,\pi)\,P_\circ(\pi)} \\ &= \frac{0.95\times 0.1}{0.95\times 0.1 + 0.02\times 0.9} = 0.84\,, \\ P(\pi\,|\,T) &= \ldots = 16\% \quad [= 1 - P(\pi\,|\,T)]\,. \end{aligned}$$

(c) **Answer: 5.3**

Detailed solution:

$$S/N \equiv \frac{P(\mu\,|\,T)}{P(\pi\,|\,T)} \;=\; \frac{P(T\,|\,\mu)\,P_\circ(\mu)}{P(T\,|\,\pi)\,P_\circ(\pi)} = \frac{0.95\times 0.1}{0.02\times 0.9} = 5.3$$

(d) **Answer: 171**

Detailed solution: The final odds, that in this example are equivalent to the S/N, can be written in terms of initial odds and Bayes Factor as

$$\frac{P(\mu\,|\,T)}{P(\pi\,|\,T)} \;=\; \frac{P(T\,|\,\mu)}{P(T\,|\,\pi)}\times\frac{P_\circ(\mu)}{P_\circ(\pi)}$$

where the first factor in the r.h.s is the so called Bayes Factor of the hypothesis μ against the hypothesis π. Now, being 95% confident on the *generic hypothesis* H, means final odds of $P(H)/P(\overline{H}) = 95/5 = 19$. In order to achieve it in the case of "muon Vs pions" and starting from initial odds $1/9 = 0.111$, we should have a Bayes Factor of

$$\frac{95/5}{10/90} = 171$$

instead of $0.95/0.02 = 47.5$.

(e) **Answer: 0.58%**

Detailed solution: To get a Bayes Factor of 171 ('BF_{wish}') with $P(T\,|\,\mu)\approx 1$ we need

$$P^{(wish)}(T\,|\,\pi) \;=\; \frac{\approx 1}{BF_{wish}} = 0.0058\,.$$

4. (a) **Answer: 45.5%**

Detailed solution:

$$\begin{aligned}\frac{P(\text{HIV}\,|\,\text{Pos})}{P(\overline{\text{HIV}}\,|\,\text{Pos})} &= \frac{P(\text{Pos}\,|\,\text{HIV})}{P(\text{Pos}\,|\,\overline{\text{HIV}})}\times\frac{P_0(\text{HIV})}{P_0(\overline{\text{HIV}})}\\ &= \frac{1}{0.002}\times\frac{0.1}{60} = 500\times\frac{1}{600} = \frac{1}{1.2} = 0.833\;\; < \mathbf{1}\,,\end{aligned}$$

from which it follows

$$P(\text{HIV}\,|\,\text{Pos}) \;=\; \frac{0.833}{1.833} = 0.455$$

(b) **Answer: 99.76%**

Detailed solution: Being independent pieces of evidence, the overall Bayes Factor is the product of the two and hence

$$500 \times 500 \times \frac{1}{600} = \frac{250000}{600} = 417$$

from which it follows a probability of $417/418 = 0.9976$.

(c) **Answer: 89.3%**

Detailed solution:

$$\begin{aligned}
\frac{P(\mathrm{HIV}\,|\,\mathrm{Pos}^{(1)}, \mathrm{Pos}^{(2)})}{P(\overline{\mathrm{HIV}}\,|\,\mathrm{Pos}^{(1)}, \mathrm{Pos}^{(2)})} &= \frac{P(\mathrm{Pos}^{(1)}, \mathrm{Pos}^{(2)}\,|\,\mathrm{HIV})}{P(\mathrm{Pos}^{(1)}, \mathrm{Pos}^{(2)}\,|\,\overline{\mathrm{HIV}})} \times \frac{P_0(\mathrm{HIV})}{P_0(\overline{\mathrm{HIV}})} \\
&= \frac{1}{P(\mathrm{Pos}^{(2)}\,|\,\mathrm{Pos}^{(1)}, \overline{\mathrm{HIV}}) \times P(\mathrm{Pos}^{(1)}\,|\,\overline{\mathrm{HIV}})} \times \frac{P_0(\mathrm{HIV})}{P_0(\overline{\mathrm{HIV}})} \\
&= \frac{1}{0.1 \times 0.002} \times \frac{0.1}{60} = 5000 \times \frac{1}{600} = 8.33
\end{aligned}$$

from which

$$P(\mathrm{HIV}\,|\,\mathrm{Pos}^{(1)}, \mathrm{Pos}^{(2)}) = \frac{8.33}{9.33} = 89.3\%.$$

5. (a) **Answer:** $f(\epsilon\,|\,x = n = 100) = 101\,\epsilon^{100}$

Detailed solution: In general

$$\begin{aligned}
f(\epsilon\,|\,x, n) &\propto f(x\,|\,\epsilon, n)\, f_0(\epsilon) \\
&\propto \frac{n!}{x!\,(n-x)!}\,\epsilon^x\,(1-\epsilon)^{n-x}\, f_0(\epsilon)
\end{aligned}$$

from which, in the special case $x = n = 100$ and $f_0(\epsilon) = k$

$$\begin{aligned}
f(\epsilon\,|\,x = n = 100) &\propto f(x = 100\,|\,\epsilon, n = 100)\, f_0(\epsilon) \\
&\propto \epsilon^{100}\,,
\end{aligned}$$

that needs to be normalized dividing it by $\int_0^1 \epsilon^{100}\, d\epsilon = 1/101$, thus getting

$$f(\epsilon\,|\,x = n = 100) = 101\,\epsilon^{100}$$

(b) **Answer: 1**

Detailed solution: $f(\epsilon\,|\,x = n = 100)$ is an increasing function of ϵ in the range $0 \le \epsilon \le 1$, so its maximum is for $\epsilon = 1$

(c) **Answer:** $101/102 \approx 0.9902$

Detailed solution:

$$\mathrm{E}[\epsilon\,|\,x = n = 100] = \int_0^1 \epsilon\, f(\epsilon\,|\,x = n = 100)\, d\epsilon = \int_0^1 101\,\epsilon^{101}\, d\epsilon = \frac{101}{102}$$

(d) **Answer:** $\epsilon_m = 1$; $\mathbf{E}[\epsilon] = 0.9902$; $\sigma(\epsilon) = 0.0097$

Detailed solution: A flat initial pdf of a continuous variable in the range $[0, 1]$ can be described by a beta with $r_i = 1$ and $s_i = 1$. The experiment updates r and s by the rule

$$\begin{aligned} r_f &= r_i + x \\ s_f &= s_i + (n - x) \end{aligned}$$

thus getting $r_f = 101$ and $s_f = 1$. Mode, expected value and standard deviation of a beta (the formula for the mode is valid only if $r \geq 1$ and $s \geq 1$):

$$\begin{aligned} \text{mode} &= \frac{r-1}{r+s-2} \\ \text{expected value} &= \frac{r}{r+s} \\ \text{standard deviation} &= \sqrt{\frac{r\,s}{(r+s-1)\,(r+s)^2}}\,, \end{aligned}$$

from which we recover the above results for ϵ_m and $\mathrm{E}[\epsilon]$, as well as $\sigma(\epsilon) = 0.0097$.

6. **Answer:** $p = 0.54 \pm 0.05$

Detailed solution: Modelling the prior pdf by a beta with expected value 0.75 and standard deviation 0.1, we can invert the formulae in order to get r and s. The approximation $r + s + 1 \approx r + s$ simplifies the calculation,[1] thus getting the inverse relations

$$\begin{aligned} s &\approx \frac{\alpha}{\sigma^2(p)\,(1+\alpha)^3} \\ r &= \alpha s\,, \end{aligned}$$

having introduced $\alpha = r/s = \frac{\mathrm{E}[p]}{1-\mathrm{E}[p]}$ just to simplify the notation.

We get then

$$\begin{aligned} r_i &\approx 14.1 \\ s_i &\approx 4.7\,, \end{aligned}$$

[1]The exact formulae are

$$\begin{aligned} r &= \mathrm{E}^2[p]\,\frac{1-\mathrm{E}[p]}{\sigma^2(p)} - \mathrm{E}[p] \\ s &= \mathrm{E}[p]\,\frac{(1-\mathrm{E}[p])^2}{\sigma^2(p)} + \mathrm{E}[p] - 1\,, \end{aligned}$$

from which we would have got $r_i = 13.3$ and $s_i = 4.4$, and hence $r_f = 63.3$ and $s_f = 54.4$, with absolutely negligible effect on the final result.

updated by the experiment to

$$\begin{aligned} r_f &\approx 14.1 + 50 = 64.1 \\ s_f &\approx 4.7 + 50 = 54.7 \end{aligned}$$

from which we get $\mathrm{E}[p\,|\,\mathrm{data}] = 0.54$ and $\sigma(p\,|\,\mathrm{data}) = 0.05$.

7. (a) **Answer:** $\mu_1 = 1.05 \pm 0.11\,\mathbf{u}$; $\mu_2 = 1.23 \pm 0.11\,\mathbf{u}$

Detailed solution: Calling μ_{R_i} the values neglecting possible correction for systematics, the offset error z, of value $0.00 \pm 0.10\,\mathrm{u}$, adds to it, thus yielding

$$\mu_i = \mu_{R_i} + z\,.$$

It follows, being μ_i and z independent

$$\begin{aligned} \mathrm{E}[\mu_i] &= \mathrm{E}[\mu_{R_i}] + \mathrm{E}[z] = \mathrm{E}[\mu_{R_i}] \\ \sigma^2\mu_i) &= \sigma^2(\mu_{R_i}) + \sigma^2(z) \end{aligned}$$

(b) **Answer:** $\rho = 0.862$

Detailed solution: Being $\mu_i = \mu_{R_i} + z$ we have, explicitly

$$\begin{aligned} \mu_1 &= \mu_{R_1} + z \\ \mu_2 &= \mu_{R_2} + z\ : \end{aligned}$$

the common term z in the linear transformation correlates the results. In fact, the covariance among μ_1 and μ_2 is given by

$$\mathrm{cov}(\mu_1, \mu_2) = \frac{\partial \mu_1}{\partial z}\frac{\partial \mu_2}{\partial z}\,\sigma^2(z) = \sigma^2(z)$$

The correlation coefficient is then

$$\begin{aligned} \rho(\mu_1, \mu_2) &= \frac{\mathrm{cov}(\mu_1, \mu_2)}{\sigma(\mu_1)\,\sigma(\mu_2)} = \frac{\sigma^2(z)}{\sigma(\mu_1)\,\sigma(\mu_2)} \\ &= \frac{(0.10\,\mathrm{u})^2}{\sqrt{0.10^2 + 0.04^2}\,\mathrm{u} \times \sqrt{0.10^2 + 0.04^2}\,\mathrm{u}} = 0.862 \end{aligned}$$

(Be careful rounding correlation coefficients close to 1: they might influence later 'propagations'!)

(c) **Answer:** $\Delta\mu = 0.18 \pm 0.06\,\mathbf{u}$

Detailed solution: The common additive contribution cancels in differences, since indeed

$$\Delta\mu = \mu_2 - \mu_1 = \mu_{R_2} - \mu_{R_1}$$

(d) **Answer:** $M = 2.28 \pm 0.21\,\mathbf{u}$

Detailed solution: Differently to the previous case we have now

$$\begin{aligned} M &= \mu_1 + \mu_2 = \mu_{R_1} + \mu_{R_2} + 2\,z \\ \sigma^2(M) &= \sigma^2(\mu_{R_1}) + \sigma^2(\mu_{R_2}) + 4\,\sigma^2(z). \end{aligned}$$