First homework assignment - Solution

Question 1)
Answer:
item a)

Tanto $Pr(X = 1|\theta = 0)$ quanto $Pr(X = 0|\theta = 1)$ são obtidas a partir do seguinte trecho:

"if you have the disease, it will correctly say so 99% of the time; if you don't have the disease, it will correctly say so 99% of the time."

Então $Pr(X = 1|\theta = 0) = 0.01$ e $Pr(X = 0|\theta = 1) = 0.01$.

$Pr(\theta = 1)$ é obtido a partir do seguinte trecho:

"But the disease in question is very rare; just one person in every 10,000 has it. This is known as your “prior probability” "

Então $Pr(\theta = 1) = 0.0001$.

item b)

Use a regra de Bayes para calcular $Pr(\theta = 1|X = 1)$

$$\begin{aligned}
Pr(\theta = 1|X = 1) &= \frac{Pr(X = 1|\theta = 1)Pr(\theta = 1)}{Pr(X = 1)} \\
&= \frac{Pr(X = 1|\theta = 1)Pr(\theta = 1)}{Pr(X = 1|\theta = 1)Pr(\theta = 1) + Pr(X = 1|\theta = 0)Pr(\theta = 0)} \\
&= \frac{0.99 \times 0.0001}{0.99 \times 0.0001 + 0.0001 \times 0.9999} \\
&\approx 0.0098039
\end{aligned}$$

Note que $Pr(\theta = 1) = 0.0001$ enquanto $Pr(\theta = 1|X = 1) = 0.0098039$. Ao receber um resultado positivo do teste a probabilidade de que a pessoa esteja de fato doente aumenta aproximadamente em 98 vezes.

item c)

O trecho "If you took this test entirely at face value, then you'd be scaring a lot of people, and sending them for intrusive, potentially dangerous medical procedures, on the back of a misdiagnosis.” destaca o componente de falsos positivos $Pr(X = 1|\theta = 0)$ $Pr(\theta = 0)$ no calculo de $Pr(X = 1)$. No exemplo apresentado,

$$Pr(X = 1|\theta = 0)Pr(\theta = 0) = 0.009999 > 0.000099 = Pr(X = 1|\theta = 1)Pr(\theta = 1)$$

levando ao "If you took this test entirely at face value, then you'd be scaring a lot of people".

item d)

Considere agora que $Pr(\theta = 1) = 0.01$

$$\begin{aligned}
Pr(\theta = 1|X = 1) &= \frac{Pr(X = 1|\theta = 1)Pr(\theta = 1)}{Pr(X = 1)} \\
&= \frac{Pr(X = 1|\theta = 1)Pr(\theta = 1)}{Pr(X = 1|\theta = 1)Pr(\theta = 1) + Pr(X = 1|\theta = 0)Pr(\theta = 0)} \\
&= \frac{0.99 \times 0.01}{0.99 \times 0.01 + 0.01 \times 0.99} \\
&= 0.5
\end{aligned}$$

Note que $Pr(\theta = 1) = 0.01$ e $Pr(\theta = 1|X = 1) = 0.5$. Similarmente ao caso do item b), um teste positivo levará a uma maior probabilidade (50 vezes maior) do que antes de observarmos $X = 1$.

Question 2)

Lembre-se do caso Normal - Normal:

$$X_1, \ldots, X_n \sim N(\theta, \sigma^2)$$

$$\theta \sim N(\mu_0, \tau_0^2)$$

então

$$\theta|\sigma^2, X_1, \ldots, X_n \sim N(\mu_n, \tau_n^2)$$

em que

$$\mu_n = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n}{\sigma^2}\bar{X}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}$$

$$\tau_n^2 = \frac{1}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}$$

Answer:

item a)

$Pr(\theta < 800|A) = pnorm(800, 900, 20) = 2.866516e - 07$

$Pr(\theta < 800|A, X = 850) = pnorm(800, 890, \sqrt{320}) = 2.437695e - 07$

$Pr(\theta < 800|B) = pnorm(800, 800, 80) = 0.5$

$Pr(\theta < 800|B, X = 850) = pnorm(800, 840, \sqrt{1280}) = 0.1317762$

Físico A é mais experiente tendo uma priori mais concentrada em torno de 900 enquanto o físico B é menos experiente tendo uma priori mais difusa. Após observar $X_1 = 850$, ambos encontram uma menor probabilide de que $\theta < 800$ do que antes de observar X_1. Note ainda que a variação da probabilidade de $\theta < 800$ antes e depois de observar $X_1 = 850$ é maior para o físico B devido a modelagem com uma priori mais difusa.

item b)

$Pr(\theta < 800|A, X_1 = 850, X_2 = 850) = pnorm(800, 883.33, \sqrt{266.66}) = 1.67064e - 07$

$Pr(\theta < 800|B, X_1 = 850, X_2 = 850) = pnorm(800, 844.44, \sqrt{711.11}) = 0.04779035$

Similarmente ao item a, após observar $X_1 = 850$ e $X_2 = 850$, ambos encontram uma menor probabilide de que $\theta < 800$ do que antes de observar X_1 ou após observar $X_1 = 850$ mas antes de observar X_2. Novamente a variação da probabilidade de $\theta < 800$ antes e depois de obser $X_2 = 850$ é maior para físico B devido a modelagem com uma priori mais difusa. Por fim, note que ao observarmos $X_i = 850, i = 1, 2$, ambos os fisicos atribuem uma probabilidade de $\theta < 800$ cada vez mais próxima a zero.