Ch.2-Ex.17: In the simulation experiment, $X_i \sim \text{i.i.d.} N(0, 1)$. Hence, $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N(0, \sqrt{\frac{1}{n}})$. These box plots in Figure 2.27 show the dispersion of $\bar{X}$, with $n=4, 16, 64, 256$ respectively.

For (a), $n=4$. Then, $Sd(\bar{X}) = \sqrt{\frac{1}{4}} = \frac{1}{2}$.

For (b), $n=16$. Then, $Sd(\bar{X}) = \sqrt{\frac{1}{16}} = \frac{1}{4}$.

For (c), $n=64$. Then, $Sd(\bar{X}) = \sqrt{\frac{1}{64}} = \frac{1}{8}$.

For (d), $n=256$. Then, $Sd(\bar{X}) = \sqrt{\frac{1}{256}} = \frac{1}{16}$.

We can see that from (a) to (b), the vertical scale decrease by 50% every time. The actual decrease degree is consistent with the formula given above.

Ch.2-Ex.18:

$$Pr(D = 1|T = 0) = \frac{Pr(D = 1 \text{ and } T = 0)}{Pr(T = 0)}$$

$$= \frac{Pr(T = 0|D = 1) Pr(D = 1)}{Pr(T = 0|D = 1) Pr(D = 1) + Pr(T = 0|D = 0) Pr(D = 0)}$$

$$= \frac{0.05 \times 0.001}{0.05 \times 0.001 + 0.95 \times 0.999}$$

$$= 0.05 \times 0.001 \times 5.27 \times 10^{-5}$$

Ch.2-Ex.19: Suppose $W = 1$ denotes B has weapons and $W = 0$ denotes B doesn’t have weapons. Similarly, let $I = 1$ denotes that inspectors find weapons and $I = 0$ denotes they don’t.

$$Pr(W = 1|I = 0) = \frac{Pr(W = 1 \text{ and } I = 0)}{Pr(I = 0)}$$

$$= \frac{Pr(I = 0|W = 1) Pr(W = 1)}{Pr(I = 0|W = 1) Pr(W = 1) + Pr(I = 0|W = 0) Pr(W = 0)}$$

$$= \frac{0.8 \times 0.2}{0.8 \times 0.8 + 1 \times 0.2}$$

$$= 0.762$$
Ch.2-Ex.20  (a) We fix $\lambda = 1$ for this part (though any positive value could have been used).

(b) R code:

```r
lambda.true=1; n=10; y=rexp(n,rate=lambda.true)
y
```

(c) The plot of $l(\lambda)$ is as follows:

\[
l(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda t_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} t_i}
\]

Since $T_i \overset{iid}{\sim} \text{Exp}(\lambda)$, the likelihood function is

\[
l(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda t_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} t_i}
\]

To find mle of $\lambda$, we can just maximize $\ln(l(\lambda))$, Notice that $\log(l(\lambda)) = n \log(\lambda) - \lambda \sum_{i=1}^{n} t_i$ and hence taking derivative, $\partial \log(l(\lambda))/\partial \lambda = n/\lambda - \sum_{i=1}^{n} t_i$. Equating it to zero yields, $\frac{n}{\lambda} = \sum_{i} t_i \Rightarrow \lambda = \frac{n}{\sum_{i} t_i} = 1/\bar{t}$. Hence, m.l.e. of $\lambda$ is $\hat{\lambda} = 1/\bar{t} = 1.76$ (this value is not unique, as it depends on the random sample we generated).

You can use R to find the $\lambda$ such that maximizes the likelihood function. Try the following R code.
lik=function(lambda){prod(dexp(y,rate=lambda))}
likelihood=function(lambda){sapply(lambda,lik)}
curve(likelihood,from=0, to=4,
xlab=expression(lambda),ylab=expression(l(lambda)))

lambda.hat=1/mean(y)

#alternatively you can find the maximum numerically
optimize(likelihood,c(0,10),maximum=T)

(d) The accuracy of mle can be measured by the LS$_{0.1}$ interval which can be approximated by $[\hat{\lambda} - \frac{2}{\lambda \sqrt{n}}, \hat{\lambda} + \frac{2}{\lambda \sqrt{n}}]$. Notice that $|\lambda - \hat{\lambda}| = |1.76 - 1| = 0.76 > 2/\lambda \sqrt{n} = 0.63$. The 95% interval $[1.13, 2.39]$ is not very accurate as the true value 1 lies outside.

abline(v=lambda.hat,col="blue")
abline(v=lambda.true,col="red")
lower=lambda.hat-2/(lambda.true*sqrt(n))
upper=lambda.hat+2/(lambda.true*sqrt(n))
c(lower,upper)
abline(v=lower,col="gray", lty=2)
abline(v=upper,col="gray", lty=2)

(e) Notice that

$$\frac{l(\lambda_1)}{l(\lambda_2)} = \left(\frac{\lambda_1}{\lambda_2}\right)^n \frac{\exp\{-\lambda_1 \sum_{i=1}^n t_i\}}{\exp\{-\lambda_2 \sum_{i=1}^n t_i\}} = \left(\frac{\lambda_1}{\lambda_2}\right)^n \exp\{-(\lambda_1 - \lambda_2) \sum_{i=1}^n t_i\}$$

We can see that the above expression depends only on $\sum_{i=1}^n t_i$ (and $n$), not on the values of individual $t_i$'s.

Ch.2-Ex.21 Coin A: fair, so Pr[H|A] = 0.5 = Pr[T|A]
Coin B: two-headed, so Pr[H|B] = 1 and Pr[T|B] = 0.
Also we are told, Pr[A] = Pr[B] = 0.5.

(a) Pr(H) = Pr(H|A) Pr(A) + Pr(H|B) Pr(B) = (0.5) * (0.5) + 1 * (0.5) = 0.75
(b) next we want find Pr[B|H]:

$$\Pr(B|H) = \frac{\Pr(B \text{ and } H)}{\Pr(H)} = \frac{\Pr(H|B) \Pr(B)}{\Pr(H)} = \frac{1 \times 0.5}{0.75} \text{ (using part (a))} = 0.6667$$
(c) next we want find $\Pr[B|T]$:

$$\Pr(B|H) = \frac{\Pr(B \text{ and } T)}{\Pr(T)} = \frac{\Pr(T|B) \Pr(B)}{\Pr(T)} = \frac{0 \times 0.5}{1 - 0.75} \quad \text{(using part (a))} = 0$$

(d) If you randomly choose a coin at the second time, $\Pr(H \text{ in 2nd}|H \text{ in 1st}) = \Pr(H \text{ in 2nd}) = \Pr(H) = 0.75$.

Ch.2-Ex.22 We are given, $\Pr[H|A] = 0.25 = 1 - \Pr[T|A]$ and $\Pr[H|B] = 0.6667 = 1 - \Pr[T|B]$ and $\Pr[A] = \Pr[B] = 0.5$.

(a) $\Pr[H] = \Pr[H|A] \times \Pr[A] + \Pr[H|B] \times \Pr[B] = (0.25) \times (0.5) + (0.6667) \times (0.5) = 0.4583$

(b) $\Pr[A|H] = \frac{\Pr[H|A] \times \Pr[A]}{\Pr[H]} = \frac{(0.25) \times (0.5)}{0.4583} = 0.2727$

$\Pr[A|T] = \frac{\Pr[T|A] \times \Pr[A]}{\Pr[T]} = \frac{(1 - 0.25) \times (0.5)}{(1 - 0.4583)} = 0.6923$

(c) If you randomly choose a coin at the second time, $\Pr(H \text{ in 2nd}|H \text{ in 1st}) = \Pr(H \text{ in 2nd}) = \Pr(H) = 0.4583$.

Ch.2-Ex.23 Suppose $M = 1$ denotes math majors and $M = 0$ denotes non-math majors. Let $X$ denotes the SAT scores. We are given $X|M = 1 \sim N(\mu = 700, \sigma = 50)$ and $X|M = 0 \sim N(\mu = 600, \sigma = 50)$ and $\Pr[M = 1] = 0.05 = 1 - \Pr[M = 0]$.

We want to find $\Pr[M = 1|X = 720]$. By Bayes theorem,

$$\Pr[M = 1|X = 720] = \frac{f(720; \mu = 700, \sigma = 50) \times \Pr[M = 1]}{f(720; \mu = 700, \sigma = 50) \times \Pr[M = 1] + f(720; \mu = 600, \sigma = 50) \times \Pr[M = 0]}$$

$$= \frac{(0.00736) \times (0.05)}{(0.00736) \times (0.05) + (0.000448) \times (0.95)} = 0.46395$$

where $f(x; \mu, \sigma) = (\sqrt{2\pi}\sigma)^{-1}\exp\left\{-(x - \mu)^2/2\sigma^2\right\}$ denotes pdf of $N(\mu, \sigma)$ i.e., a normal distribution with mean $\mu$ and standard deviation $\sigma$ Use the following R code:

$$\text{dnorm}(720, 700, 50) \times 0.05 / (\text{dnorm}(720, 700, 50) \times 0.05 + \text{dnorm}(720, 600, 50) \times 0.95)$$