1.(a) Straightforward calculations:

$$\begin{aligned}
\mathbb{P}(X=1,Y=2) &= \mathbb{P}(X=1)\mathbb{P}(Y=2) \text{ [by independence]}\\
&= \left[\binom{10}{1}0.1^1(1-0.1)^{10-1}\right]\left[\binom{20}{2}0.1^2(1-0.1)^{20-2}\right]\\
&= 0.1105.
\end{aligned}$$

(b) Let $R = X+Y$. Theorem 2.1 from the course book implies that $R \sim \text{Binomial}(m+n, p)$. Hence, $R \sim \text{Binomial}(30, 0.1)$. Using the pmf for the Binomial distribution, we get:

$$\begin{aligned}
\mathbb{P}(X+Y=3) &= \mathbb{P}(R=3)\\
&= \binom{30}{3}0.1^3(1-0.1)^{30-3}\\
&= 0.2361.
\end{aligned}$$

(c) Theorem from the Section 2.4.1 of the course book (see "Conditioning") implies that $X|R=3 \sim \text{Hypergeometric}(3, 10, 30)$.

(d) We apply the formula for the expectation for the Hypergeometric distribution (see the course book, page 15):

$$\mathbb{E}(X|R=3) = \frac{3\times 10}{30} = 1.$$

(e) According to the Theorem from the Section 2.4.1 of the course book (see "Conditioning"), the conditional pmf of X given that $R=3$ does not depend on p. Hence, $\mathbb{E}(X|R=3) = 1$ disregarding the value of p.

Comment: We know that $R = X+Y = 3$. This means that we have a population with 3 diseased people. It is evident that the number of the healthy people in the population is $m+n-3 = 30-3 = 27$. Once we have this information, we are not concerned about the value of p that originally generated the population.

2.(a) Under the hypothesis that the coin is fair in each experiment, we have:
$X_1 \sim \text{Binomial}(104, 0.5)$,
$X_2 \sim \text{Binomial}(84, 0.5)$,
$X_3 \sim \text{Binomial}(86, 0.5)$.

(b) For Experiment$_1$: $X_1 \sim \text{Binomial}(104, p)$ and

$$\begin{aligned}
H_0 &: \quad p = 0.5,\\
H_1 &: \quad p \neq 0.5 \text{ (two} - \text{sided)}.
\end{aligned}$$

(c) The results for all three experiments are presented below:
Experiment$_1$:

$$\begin{aligned}
p\text{−value} &= 2\times\mathbb{P}(X_1 \geq 81)\\
&= 2(1-\mathbb{P}(X_1 \leq 80))\\
&= 2*(1-\text{pbinom}(80, 104, 0.5)) \text{ [R command]}\\
&= 9.3\times 10^{-9}.
\end{aligned}$$

Experiment$_2$:

$$\begin{aligned}
p\text{−value} &= 2 \times \mathbb{P}(X_1 \geq 68) \\
&= 2(1 - \mathbb{P}(X_1 \leq 67)) \\
&= 2 * (1 - \text{pbinom}(67, 84, 0.5)) \text{ [R command]} \\
&= 8.5 \times 10^{-9}.
\end{aligned}$$

Experiment$_3$:

$$\begin{aligned}
p\text{−value} &= 2 \times \mathbb{P}(X_1 \geq 57) \\
&= 2(1 - \mathbb{P}(X_1 \leq 56)) \\
&= 2 \times (1 - \text{pbinom}(56, 86, 0.5)) \text{ [R command]} \\
&= 0.0034.
\end{aligned}$$

(d) In Experiment$_3$, we have the weakest evidence against the null hypothesis.

3.(a) When $\lambda = 1$, $X \sim \text{Geometric}\left(\frac{1}{2}\right)$. Therefore, for $x = 0, 1, 2, \ldots$, we have:

$$\mathbb{P}(X = x) = \left(1 - \frac{1}{2}\right)^x \frac{1}{2} = \left(\frac{1}{2}\right)^{x+1}.$$

Applying the definition of the cumulative distribution function, we get:

$$\begin{aligned}
F_X(1.99) &= \mathbb{P}(X \leq 1.99) \\
&= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) \\
&= \left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^2 \\
&= \frac{3}{4}.
\end{aligned}$$

Furthermore, we have:

$$\begin{aligned}
\mathbb{P}(X > 2) &= 1 - \mathbb{P}(X \leq 2) \\
&= 1 - F_X(2) \\
&= 1 - [\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)] \\
&= 1 - \left[\left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3\right] \\
&= 1 - \frac{7}{8} \\
&= \frac{1}{8}.
\end{aligned}$$

(b) Let k and t be nonnegative integers. For ease of writing, we define $p = \dfrac{1}{\lambda + 1}$. It is obvious that $0 < p < 1$. Using the formula for the conditional probabilities, we get:

$\mathbb{P}(X = k + t | X \geq k)$

$$
\begin{aligned}
&= \frac{\mathbb{P}\left(\{X = k+t\} \cap \{X \geq k\}\right)}{\mathbb{P}\left(\{X \geq k\}\right)} \\
&= \frac{\mathbb{P}(X = k+t)}{\mathbb{P}(X \geq k)} \text{ [because } t \geq 0] \\
&= \frac{\mathbb{P}(X = k+t)}{\mathbb{P}(X = k) + \mathbb{P}(X > k)} \\
&= \frac{p(1-p)^{k+t}}{p(1-p)^k + (1-p)^{k+1}} \text{ [pmf and cdf Geometric distrib., page 15 course book]} \\
&= \frac{p(1-p)^{k+t}}{(1-p)^k(p+1-p)} \\
&= \frac{p(1-p)^{k+t}}{(1-p)^k} \\
&= p(1-p)^t \\
&= \mathbb{P}(X = t) \text{ [pmf Geometric distribution]}
\end{aligned}
$$

(c) For $0 < \lambda < \infty$, the likelihood function is:

$$
\begin{aligned}
&L(\lambda; x_1, x_2, \ldots, x_n) \\
&\quad = \mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) \\
&\quad = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2)\cdots\mathbb{P}(X_n = x_n) \text{ [by independence]} \\
&\quad = \left[\left(1 - \frac{1}{\lambda+1}\right)^{x_1}\left(\frac{1}{\lambda+1}\right)\right]\cdots\left[\left(1 - \frac{1}{\lambda+1}\right)^{x_n}\left(\frac{1}{\lambda+1}\right)\right] \text{ [pmf Geometric distrib.]} \\
&\quad = \left(\frac{\lambda}{\lambda+1}\right)^{x_1+x_2+\cdots+x_n}\left(\frac{1}{\lambda+1}\right)^n \\
&\quad = \lambda^{x_1+x_2+\cdots+x_n}(\lambda+1)^{-(x_1+x_2+\cdots+x_n)-n}.
\end{aligned}
$$

As $\bar{x} = \dfrac{x_1 + x_2 + \cdots + x_n}{n}$, we get

$$
L(\lambda; x_1, x_2, \ldots, x_n) = \lambda^{n\bar{x}}(\lambda+1)^{-n\bar{x}-n}.
$$

Hence, for $0 < \lambda < \infty$, the log-likelihood function is:

$$
\begin{aligned}
\log L(\lambda; x_1, x_2, \ldots, x_n) &= \log\left(\lambda^{n\bar{x}}(\lambda+1)^{-n\bar{x}-n}\right) \\
&= n\bar{x}\log(\lambda) - (n\bar{x} + n)\log(\lambda+1) \\
&= -n(\bar{x}+1)\log(\lambda+1) + n\bar{x}\log(\lambda).
\end{aligned}
$$

(d) Based on the result from part (c), we have:

$$
\begin{aligned}
\frac{d\log L}{d\lambda} &= n\left(-\frac{\bar{x}+1}{\lambda+1} + \frac{\bar{x}}{\lambda}\right) \\
&= n\frac{-\lambda(\bar{x}+1) + (\lambda+1)\bar{x}}{\lambda(\lambda+1)} \\
&= n\frac{-\lambda\bar{x} - \lambda + \lambda\bar{x} + \bar{x}}{\lambda(\lambda+1)} \\
&= n\frac{-\lambda + \bar{x}}{\lambda(\lambda+1)},
\end{aligned}
$$

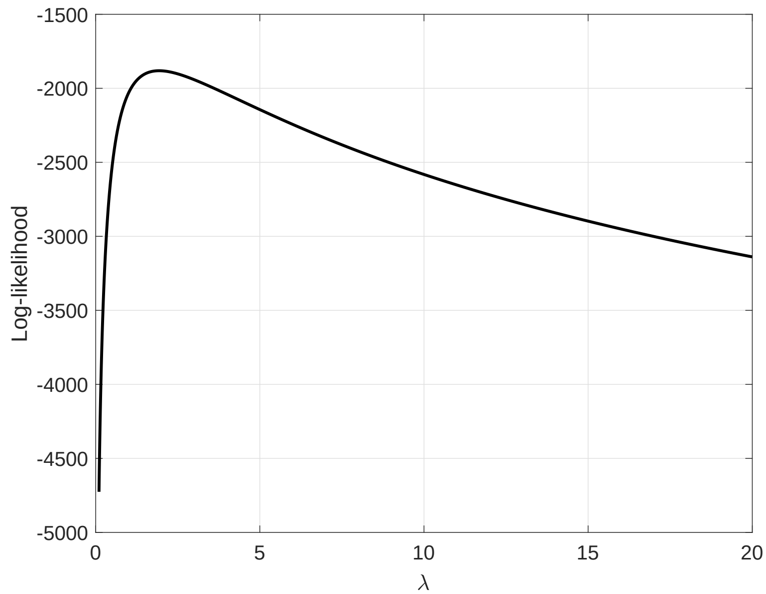

Figure 1: Plot for question 3, part (d). Note that $\bar{x} = 1.93$.

$$\frac{d \log L}{d\lambda} = 0 \quad \Rightarrow \quad \lambda = \bar{x}.$$

As the graph presented in Figure 1 confirms that there is a unique maximum for the log-likelihood function, we conclude that the maximum likelihood estimate is $\hat{\lambda} = \bar{x}$.