1.

(a)

Denote Narvey Dent picks up a double-headed coin as event "DH", a double-tailed coin as event "DT", and a normal one as "N". Denote the lower face of the coin is a head as event "LH". According to Bayes' theorem,

$$\begin{aligned}P[\mathrm{LH}] =& P[\mathrm{LH}|\mathrm{DH}]P[\mathrm{DH}] + P[\mathrm{LH}|\mathrm{DT}]P[\mathrm{DT}] + P[\mathrm{LH}|\mathrm{N}]P[\mathrm{N}]\\ =&1\cdot\frac{2}{5} + 0\cdot\frac{1}{5} + \frac{1}{2}\cdot\frac{2}{5}\\ =&\frac{3}{5}\end{aligned}$$

So the probability is $\frac{3}{5}$.

(b)

Denote the top face of the coin showing heads as "TH". Then,

$$\begin{aligned}P[\mathrm{LH}|\mathrm{TH}] =&\frac{P[\mathrm{TH}\cap\mathrm{LH}]}{P[\mathrm{TH}]} = \frac{P[\mathrm{DH}]}{P[\mathrm{TH}|\mathrm{DH}]P[\mathrm{DH}] + P[\mathrm{TH}|\mathrm{DT}]P[\mathrm{DT}] + P[\mathrm{TH}|\mathrm{N}]P[\mathrm{N}]}\\ =&\frac{\frac{2}{5}}{1\cdot\frac{2}{5} + 0 + \frac{1}{2}\cdot\frac{2}{5}}\\ =&\frac{2}{3}\end{aligned}$$

So the probability is $\frac{2}{3}$.

(c)

After first time trial, Harvey knows that this coin cannot be a double-tailed one. Denote last time the top face of the coin showing heads as "LTH", then $P[DT|LTH] = 0$. According to 1.(b), $P[\mathrm{DH}|\mathrm{LTH}] = \frac{2}{3}$, $P[\mathrm{N}|\mathrm{LTH}] = \frac{1}{3}$. So

$$\begin{aligned}P[\mathrm{LH}|\mathrm{TH},\ \mathrm{LTH}] =&\frac{P[\mathrm{TH}\cap\mathrm{LH}\cap\mathrm{LTH}]}{P[\mathrm{TH}\cap\mathrm{LTH}]} = \frac{P[\mathrm{DH}\cap\mathrm{LTH}]}{P[\mathrm{TH}\cap\mathrm{LTH}]}\\ =&\frac{P[\mathrm{DH}|\mathrm{LTH}]}{P[\mathrm{TH}|\mathrm{DH},\mathrm{LTH}]P[\mathrm{DH}|\mathrm{LTH}] + P[\mathrm{TH}|\mathrm{N},\mathrm{LTH}]P[\mathrm{N}|\mathrm{LTH}]}\\ =&\frac{\frac{2}{3}}{1\cdot\frac{2}{3} + \frac{1}{2}\cdot\frac{1}{3}}\\ =&\frac{4}{5}\end{aligned}$$

So the probability is $\frac{4}{5}$.

2.

Denote X_k as the number of heads' appearing out of k number of trials, with the assumption of independence, then we have

$$X_k \sim Binomial(k;p)$$

and the likelihood function

$$\mathcal{L}(p;x) = f_{X|P}(x|p) = \frac{k!}{x!(k-x)!}p^x(1-p)^{k-x}$$

Then

$$f_{P|x}(p) = \frac{f_{XP}(x,p)}{f_X(x)} = \frac{f_{X|P}(x|p)f_P(p)}{f_X(x)} \propto f_{X|P}(x|p)f_P(p)$$

Given that $X_{10} < 3$, $P \sim$Beta(4,4), we get that

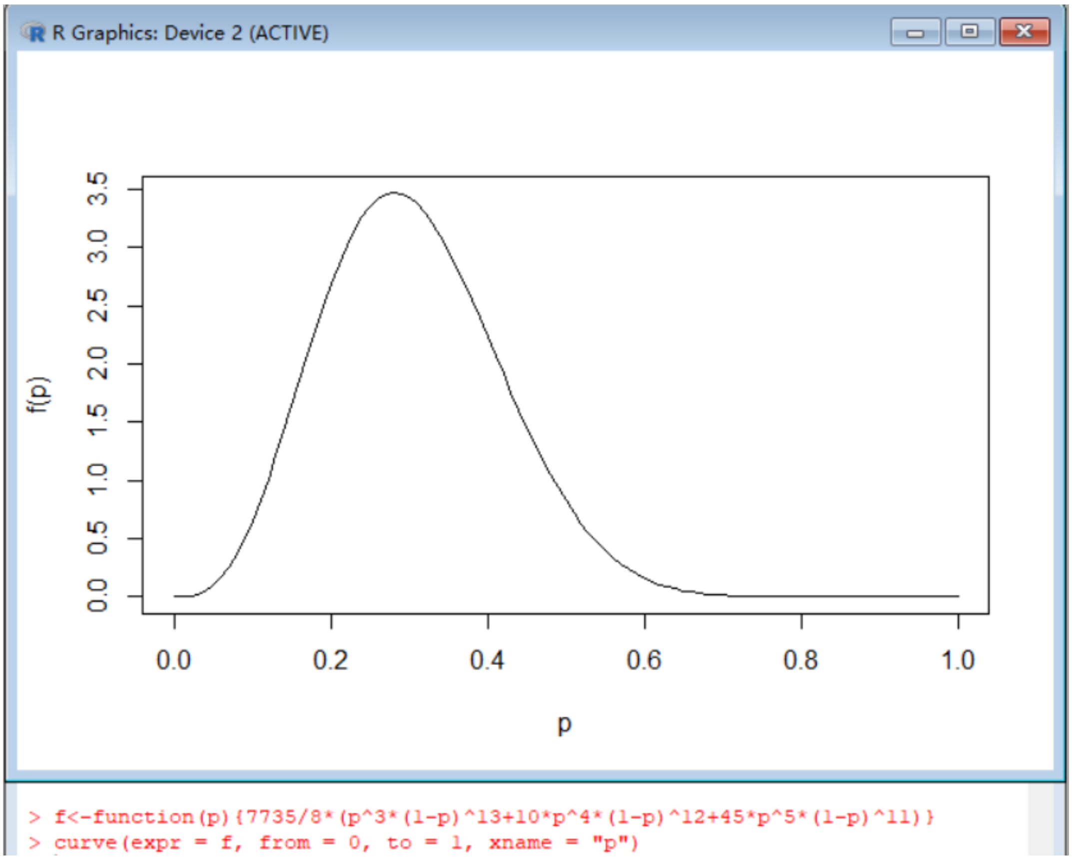
$$\begin{aligned}
f_{P|x<3}(p) =& A \cdot \Big(f_{X|P}(0|p) + f_{X|P}(1|p) + f_{X|P}(2|p)\Big) f_P(p) \\
=& A \cdot \Big(\frac{10!}{0!(10-0)!}p^0(1-p)^{(10-0)} + \frac{10!}{1!(10-1)!}p^1(1-p)^{(10-1)} \\
&+ \frac{10!}{2!(10-2)!}p^2(1-p)^{(10-2)}\Big)\frac{\Gamma(4+4)}{\Gamma(4)\Gamma(4)}p^{4-1}(1-p)^{(4-1)} \\
=& A \cdot \Big(p^3(1-p)^{13} + 10p^4(1-p)^{12} + 45p^5(1-p)^{11}\Big)
\end{aligned}$$

where A is a scale constant. Since

$$\begin{aligned}
1 =& \int_0^1 f_{P|x<3}(p)\ dp \\
=& A \cdot \Big(\frac{\Gamma(4)\Gamma(14)}{\Gamma(18)} + 10 \cdot \frac{\Gamma(5)\Gamma(13)}{\Gamma(18)} + 45 \cdot \frac{\Gamma(6)\Gamma(12)}{\Gamma(18)}\Big)
\end{aligned}$$

the posterior density for p is

$$f_{P|x}(p) = \frac{7735}{8}(p^3(1-p)^{13} + 10p^4(1-p)^{12} + 45p^5(1-p)^{11})$$



3.

Let X_1 denote the number of people choose "F" is one test, and X_2 denote the number of people choose "J". Then $X_1 + X_2 = 50$ and the probability of choosing either of items is the binomial distribution. Assume that these people would like to choose F with probability p and J with probability $1-p$. Then

$$f_{X_1X_2}(x_1, x_2) = \frac{50!}{x_1!x_2!}p^{x_1}(1-p)^{x_2} = \frac{50!}{y!(50-y)!}p^y(1-p)^{50-y} =: f_Y(y)$$

Given this binomial distribution, the likelihood function for parameter p is

$$\mathcal{L}(p; y) = f_{Y_1|p}(y_1|p)f_{Y_2|p}(y_2|p) = \frac{50!50!}{y_1!(50-y_1)!y_2!(50-y_2)!}p^{y_1+y_2}(1-p)^{50-y_1+50-y_2}$$

Assume a uniform prior $f_P(p) = \begin{cases} 1 & ,0 < p < 1 \\ 0 & ,\text{otherwise} \end{cases}$, then the posterior of p is

$$f_{P|y_1,y_2}(p) = A \cdot p^{y_1+y_2}(1-p)^{100-(y_1+y_2)} f_P(p) = \begin{cases} A \cdot p^{y_1+y_2}(1-p)^{100-(y_1+y_2)} & ,0 \leqslant p \leqslant 1 \\ 0 & ,\text{otherwise} \end{cases}$$

where A is a scale constant. Since $y_1 = 40$, $y_2 = 15$, and

$$1 = \int_{-\infty}^{\infty} f_P(p)\ dp = \int_0^1 p^{55}(1-p)^{45}\ dp$$
$$= \frac{\Gamma(56)\Gamma(46)}{\Gamma(102)}$$

$A = \dfrac{\Gamma(102)}{\Gamma(56)\Gamma(46)}$. So the posterior of p is

$$f_{P|y_1=40,y_2=15}(p) \sim Beta(56, 46)$$

R Console

```
> l = qbeta(0.025, 56, 46)
> u = qbeta(0.975, 56, 46)
> cat(paste("(",l, sep = ""), paste(u, ")", sep =""), "\n", sep =",")
(0.452219218656169,0.643998401885215),
> l = qbeta(0.05, 56, 46)
> u = qbeta(0.95, 56, 46)
> cat(paste("(",l, sep = ""), paste(u, ")", sep =""), "\n", sep =",")
(0.467773805995984,0.629166818512161),
> l = qbeta(0.1, 56, 46)
> u = qbeta(0.9, 56, 46)
> cat(paste("(",l, sep = ""), paste(u, ")", sep =""), "\n", sep =",")
(0.485774306714497,0.611847459929405),
> l = qbeta(0.15, 56, 46)
> u = qbeta(0.85, 56, 46)
> cat(paste("(",l, sep = ""), paste(u, ")", sep =""), "\n", sep =",")
(0.497943024192254,0.60004452838071),
> l = qbeta(0.2, 56, 46)
> u = qbeta(0.8, 56, 46)
> cat(paste("(",l, sep = ""), paste(u, ")", sep =""), "\n", sep =",")
(0.507619640681988,0.590604106623058),
```

Using R we see that p will fall into $(0.508, 0.591)$ with probability 60%, $(0.486, 0.612)$ with probability 80%, and $(0.452, 0.644)$ with probability 95%. And the posterior mean is given by

$$\begin{aligned}
\hat{p} =& \mathbb{E}[P] = \int_{-\infty}^{\infty} p f_P(p)\ dp = A \cdot \int_0^1 p p^{55}(1-p)^{45}\ dp \\
=& A \cdot \int_0^1 p^{56}(1-p)^{45}\ dp = \frac{\Gamma(102)}{\Gamma(56)\Gamma(46)} \frac{\Gamma(57)\Gamma(46)}{\Gamma(103)} \\
=& \frac{56}{102} > \frac{1}{2}
\end{aligned}$$

To sum up, it seems that people are biased toward F.

4.

(a)

Given the geometric distribution with PDF $f_{X_i|p}(x_i|p) = p(1-p)^{x_i-1}$, the likelihood function for parameter p is

$$\mathcal{L}(p;x) = \prod_{i=1}^{n} f_{X_i|p}(x_i|p) = p^n(1-p)^{\sum\limits_{i=1}^{n} x_i - n}$$

and we can obtain that

$$\frac{d\,\mathcal{L}(p;x)}{dp} = np^{n-1}(1-p)^{\sum\limits_{i=1}^{n} x_i - n} - (\sum_{i=1}^{n} x_i - n)p^n(1-p)^{\sum\limits_{i=1}^{n} x_i - n - 1}$$

So

$$\frac{d\,\mathcal{L}(p;x)}{dp} = 0 \Leftrightarrow n(1-p) = (\sum_{i=1}^{n} x_i - n)p \Leftrightarrow p = \frac{n}{\sum\limits_{i=1}^{n} x_i}$$

Since

$$\lim_{p \searrow 0+} \mathcal{L}(p;x) \to 0 < \mathcal{L}(\frac{1}{\bar{x}};x)$$

$\mathcal{L}(p;x_i)$ reaches its maximum at $p = \frac{1}{\bar{x}}$, i.e. $\widetilde{p} = \frac{1}{\bar{x}}$.

(b)

For $p \in (0, \frac{1}{2})$, $p = \varphi(q) = \frac{1-\sqrt{1-4q}}{2} \Rightarrow \frac{d\,\varphi(q)}{d\,q} = \frac{1}{\sqrt{1-4q}} > 0$. Then

$$\begin{aligned}\frac{d\,\mathcal{L}(\varphi(q);x)}{dq} = 0 &\Leftrightarrow \frac{d\,\mathcal{L}(\varphi(q);x)}{d\varphi(q)}\frac{d\,\varphi(q)}{d\,q} = 0 \\ &\Leftrightarrow \frac{d\,\mathcal{L}(\varphi(q);x)}{d\varphi(q)} = 0 \\ &\Leftrightarrow \varphi(q) = \frac{n}{\sum\limits_{i=1}^{n} x_i} = \frac{1}{\bar{x}}\end{aligned}$$

which is valid for $\bar{x} > 2$, and $q = \frac{\bar{x}-1}{\bar{x}^2}$, $\mathcal{L}(\varphi(q);x) = \frac{1}{\bar{x}^n}(1-\frac{1}{\bar{x}})^{n\bar{x}-n} = \frac{(\bar{x}-1)^{n\bar{x}-n}}{\bar{x}^{n\bar{x}}}$.

Similarly, for $p \in (\frac{1}{2}, 1)$, $p = \varphi(q) = \frac{1+\sqrt{1-4q}}{2} \Rightarrow \frac{d\,\varphi(q)}{d\,q} = -\frac{1}{\sqrt{1-4q}} < 0$, then

$$\varphi(q) = \frac{1}{\bar{x}}$$

which is valid for $1 < \bar{x} < 2$ and $q = \frac{\bar{x}-1}{\bar{x}^2}$, $\mathcal{L}(\varphi(q);x) = \frac{(\bar{x}-1)^{n\bar{x}-n}}{\bar{x}^{n\bar{x}}}$.

For $q = \frac{1}{4}$, we have $p = \frac{1}{2}$, $\varphi(q) = \frac{1}{\bar{x}}$ which is valid when $\bar{x} = 2$.

Finally, if $\bar{x} = 1$, then $x_i = 1$ and $\mathcal{L}(p;x) = p^n$ reaches maximum when $p \to 1$ and therefore $q \to 0$. So $\hat{q} = \frac{\bar{x}-1}{\bar{x}^2}$.

(c)

Given the uniform prior $f_P(p) = \begin{cases} 1 & ,0 < p < 1 \\ 0 & ,\text{otherwise} \end{cases}$, and the likelihood for parameter p,

$$\mathcal{L}(p; x) = \prod_{i=1}^{n} f_{X_i|p}(x_i|p) = p^n(1-p)^{\sum_{i=1}^{n} x_i - n} = (p(1-p)^{\bar{x}-1})^n$$

then the posterior of p is

$$f_{P|\bar{x}}(p) = A \cdot (p(1-p)^{\bar{x}-1})^n f_P(p) = \begin{cases} A \cdot (p(1-p)^{\bar{x}-1})^n & ,0 < p < 1 \\ 0 & ,\text{otherwise} \end{cases}$$

where A is a scale constant. Since

$$\begin{aligned} 1 = \int_{-\infty}^{\infty} f_{P|\bar{x}}(p)\ dp &= \int_0^1 (p(1-p)^{\bar{x}-1})^n\ dp \\ &= \int_0^1 p^{n+1-1}(1-p)^{n\bar{x}-n+1-1}\ dp \\ &= \frac{\Gamma(n+1)\Gamma(n\bar{x}-n+1)}{\Gamma(n\bar{x}+2)} \end{aligned}$$

$A = \dfrac{\Gamma(n\bar{x}+2)}{\Gamma(n+1)\Gamma(n\bar{x}-n+1)}$. Then the posterior mean is given by

$$\begin{aligned} \hat{p} =& \mathbb{E}[P] = \int_{-\infty}^{\infty} p f_{P|\bar{x}}(p)\ dp = A \cdot \int_0^1 p(p(1-p)^{\bar{x}-1})^n\ dp \\ =& A \cdot \int_0^1 p^{n+2-1}(1-p)^{n\bar{x}-n+1-1}\ dp = \frac{\Gamma(n\bar{x}+2)}{\Gamma(n+1)\Gamma(n\bar{x}-n+1)} \frac{\Gamma(n+2)\Gamma(n\bar{x}-n+1)}{\Gamma(n\bar{x}+3)} \\ =& \frac{n+1}{n\bar{x}+2} \end{aligned}$$

So $\hat{p} = \dfrac{n+1}{n\bar{x}+2}$.