Question 1

Let F be case in which a fair coin was chosen, and F' the case for unfair coin.

T is the described scenario in which 9 out of 10 tosses were "heads". Then:

$$P(\bar{F}|T) = \frac{P(T|\bar{F})P(\bar{F})}{P(T)} = \frac{P(T|\bar{F})P(\bar{F})}{P(T|\bar{F})P(\bar{F}) + P(T|F)P(F)}$$
$$= \frac{\binom{10}{9} \cdot 0.9^9 \cdot 0.1 \cdot \frac{1}{1000}}{\binom{10}{9} \cdot 0.9^9 \cdot 0.1 \cdot \frac{1}{1000} + \binom{10}{9} \cdot 0.5^9 \cdot 0.5 \cdot \frac{999}{1000}} = \cdots = 3.8\%$$

Question 2

Let N be the random variable for the number of children in a family.

One can easily see that N~Geo(0.5) because we stop with the birth of the first son.

Each family has exactly 1 son therefor each family has N-1 daughters. From expected value of a geometric random value and the linearity of it one can deduct that the expected value of daughters in a family is 1.

Question 3.

a) We will find MLE for $\theta = (n, p)$:

We have $x_i \sim B(n, p)$ for each {x1, x2, …,xn} therefor:

$$L(\theta|D) = \prod_{i=1}^{n} \Pr(x_i|\theta) = \prod_{i=1}^{n} \binom{n}{x_i} p^{x_i}(1-p)^{n-x_1}$$

$$\Rightarrow l(\theta|D) = \sum_{i=1}^{n} \ln(\binom{n}{x_i} p^{x_i}(1-p)^{n-x_1}) = \cdots$$

$$= \sum_{i=1}^{n} ln\binom{n}{x_i} + \sum_{i=1}^{n} x_i \ln(p) + \sum_{i=1}^{n} (n - x_i) \ln(1-p)$$

One can see that the leading constant does not affect the value of $\bar{p}$ that maximizes the likelihood, so we just ignore it:

$$\Rightarrow l(\theta|D) = \sum_{i=1}^{n} x_i \ln(p) + \sum_{i=1}^{n} (n - x_i) \ln(1-p)$$

Let us now derive this function and find for which values of p it is equal to zero:

$$\frac{\partial}{\partial p} l(\theta|D) = \frac{1}{p}\sum_{i=1}^{n} x_i - \frac{1}{1-p}\sum_{i=1}^{n} (n - x_i) = 0$$

$$\Rightarrow \bar{p} = \sum_{i=1}^{n} x_i$$

b) We will find MLE for $\theta = (\mu, \sigma^2)$:

We have $x_i \sim N(\mu, \sigma^2)$ for each {x1, x2, …,xn} therefor:

$$L(\theta|D) = \prod_{i=1}^{n} \Pr(x_i|\theta) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp(-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_i - \mu)^2)$$

$$\Rightarrow l(\theta|D) = -\frac{n}{2} ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_i - \mu)^2$$

Let us now derive this function and find for which values of $\theta = (\mu, \sigma^2)$ it is equal to zero:

(With a lot of help from wolfram-alpha we get)

$$\hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{j=1}^{n}(x_j - \hat{\mu})^2$$

c) We will find MLE for $\theta = (\lambda)$:

We have $x_i \sim Pois(\mu, \sigma^2)$ for each {x_1, x_2, …,x_n} therefor:

$$L(\theta|D) = \prod_{i=1}^{n} \Pr(x_i|\theta) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i}$$

$$\Rightarrow l(\theta|D) = -n\lambda - \sum_{i=1}^{n} x_i \ln(\lambda) - \ln(\prod_{i=1}^{n} x_i)$$

Let us now derive this function and find for which values of $\theta = (\mu, \sigma^2)$ it is equal to zero:

(With a lot of help from wolfram-alpha we get)

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Question 4.

Pearson's correlation coefficient is simply this ratio:

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

We will prove that $0 \le |\rho| \le 1$ and therefor $-1 \le \rho \le 1$.

The variance is non-negative for both X and Y by definition and so $\sqrt{Var(X)Var(Y)} \ge 0$. If X and Y are independent then $\rho = 0$, because:

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$$

Using *Cauchy-Schwarz inequality* we get:

$$|Cov(X,Y)|^2 \le Var(X)Var(Y) \Rightarrow Cov(X,Y) \le \sqrt{Var(X)Var(Y)}$$

$$\Rightarrow |\rho| = \left|\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}\right| \le \frac{\sqrt{Var(X)Var(Y)}}{\sqrt{Var(X)Var(Y)}} = 1$$

Thus we get that $0 \le |\rho| \le 1$ and therefor $-1 \le \rho \le 1$.