1. (a) Let X be r.v. for child's birth rank.
F be family number.

From LoTP, $P(X=1) = \sum_{i=1}^{3} P(X=1 | F=i)\, P(F=i)$

$$= \frac{1}{3}\left(P(X=1|F=1) + P(X=1|F=2) + P(X=1|F=3) \right)$$

$$= \frac{1}{3}\left(1 + \frac{1}{2} + \frac{1}{3}\right) = \frac{11}{18}$$

Simillarly, $P(X=2) = \frac{1}{3}\left(0 + \frac{1}{2} + \frac{1}{3}\right) = \frac{5}{18}$

$P(X=3) = \frac{1}{3}\left(0 + 0 + \frac{1}{3}\right) = \frac{1}{9}$

$$EX = \sum_{x=1}^{3} x\, P(X=x) = \frac{11}{18} + \frac{10}{18} + \frac{3}{9} = \frac{27}{18} = \frac{3}{2} = 1.5$$

$$\mathrm{Var}(X) = EX^2 - (EX)^2$$

From Lotus, $EX^2 = \frac{11}{18} + \frac{20}{18} + 1 = \frac{49}{18}$

$$\mathrm{Var}(X) = EX^2 - (EX)^2 = \frac{49}{18} - \frac{9}{4} = \frac{98-81}{36} = \frac{17}{36} \simeq 0.47$$

(b) $P(X=1) = \dfrac{\#\text{ Child with rank1}}{\#\text{ children}} = \frac{100}{190} = \frac{10}{19}$

Simillarly, $P(X=2) = \frac{70}{190} = \frac{7}{19}$

$P(X=3) = \frac{20}{190} = \frac{2}{19}$

$$EX = \frac{10}{19} + \frac{14}{19} + \frac{6}{19} = \frac{30}{19} \simeq 1.58$$

$$\mathrm{Var}(X) = EX^2 - (EX)^2 = \frac{10}{19} + \frac{28}{19} + \frac{18}{19} - \left(\frac{30}{19}\right)^2$$

$$= \frac{56}{19} - \frac{900}{19^2} = \frac{164}{361} \simeq 0.45$$

2. (a) Let X be r.v. for population of a city.

There are 10 cities. Let city Ids are $1, 2, \ldots 10$

$$\Rightarrow \quad EX = \sum_{i=1}^{10} (\text{Population of city } i) \times P(\text{city} = i)$$

from symmetry, $P(\text{city} = i) = \frac{1}{10}$

$$\Rightarrow \quad EX = \frac{1}{10} \sum_{i=1}^{10} (\text{Population of city } i)$$

$$= \frac{\text{total Population}}{10} = \frac{20}{10} = 2 \text{ million}$$

(b) Let P_i = population of city i

$$\text{Var}(X) = EX^2 - (EX)^2$$

From LOTUS, $EX^2 = \sum_{i=1}^{10} P_i^2 \times P(\text{city} = i)$

$$= \frac{1}{10} \sum_{i=1}^{10} P_i^2$$

Hence the above sum cannot be calculated unless we know population of individual cities.

(c) Let R be region

$$\Rightarrow \quad P(R = i) = \frac{1}{4} \quad \forall i = 1, 2, 3, 4$$

Let c be a city

$\forall\ i = 1, 2, 3, 4$

$\Rightarrow \quad P(c = j \mid R = i) = \dfrac{1}{\#\text{cities in } R = i}$

Let P_{ij} be population of city j in $R = i$

Let X be r.v. for population of a chosen city.

$$\Rightarrow \quad EX = \sum_{i=1}^{4} \sum P(c=j \mid R=i)\, P(R=i)\, P_{ij}$$

$$= \frac{1}{4} \left(\sum_{i=1}^{4} \sum P(c=j \mid R=i)\, P_{ij} \right)$$

$$= \frac{1}{4} \left(\frac{1}{4} \times 3 + \frac{1}{3} \times 4 + \frac{1}{2} \times 5 + \frac{1}{1} \times 8 \right)$$

$$= \frac{1}{4} \times \frac{9 + 16 + 30 + 96}{12} = \frac{151}{48} \simeq 3.15$$

(d) (c) is sum of population per region per ciy for all four regions. And for one the regions (west) this amount equals (a), Hence sum is more than (a).

3.(a) For any $k > 100$, $P(Y=k)$ will have a positive non-zero probability and for such k, $P(X \geq k) = 0$ because support for X is $[0, 100]$. Hence not possible.

(b) We can define X any Y in such a way that if $(Y=k)$ ~~then~~ occurs implies $(X \geq k)$ occurs. E.g. Let a random number generator generates number between 1 and 10. So for X success is if number generated is $[1, 9]$ and for Y success is if number generated is $[1, 5]$ in any i.i.d Bernoulli trials. ~~So, $P(X \geq k, Y \leq k)$~~

~~So, in this case $P(X \geq k \cap Y \leq k)$~~

So, in this case $P(X \geq k, Y = k) = P(Y = k)$

Hence, $P(X \geq k \mid Y = k) = \dfrac{P(Y=k)}{P(Y=k)} = 1$

Hence this possible.

(c) Not possible, because for $Y = 0$

$$P(X \leq 0 \mid Y = 0) = P(X = 0 \mid Y = 0)$$

$$= \frac{P(X=0, Y=0)}{P(Y=0)}$$

Now, $P(X=0) < P(Y=0)$

So, $P(X=0, Y=0)$ can be at most $P(X=0)$

$$\Rightarrow \frac{P(X=0, Y=0)}{P(Y=0)} \leq \frac{P(X=0)}{P(Y=0)} < 1$$

4. (a) Let Z be r.v. for ~~first~~ first time they succeded simultaneously.

In any trial both of them will succeed with Probability P_1P_2

So, $Z \sim FS(P_1P_2)$ or $Z \sim Geom(P_1P_2)$

And, $EZ = \frac{1}{P_1P_2}$

(b) Let W be r.v. for first time atleast one of them succeed.

Probability that atleast one of them will succeed $= P(\text{Nick} \cup \text{Penny}) = P_1 + P_2 - P_1P_2$

$\Rightarrow W \sim FS(P_1 + P_2 - P_1P_2)$ or $W \sim Geom(P_1+P_2-P_1P_2)$

$$EW = \frac{1}{P_1 + P_2 - P_1P_2}$$

(c) Let $P_1 = P_2 = p$

$$P(X=Y) = \sum_{k=1}^{\infty} P(X=k \mid Y=k)\,P(Y=k)$$

$$= \sum_{k=1}^{\infty} P(X=k, Y=k)$$

$$= \sum_{k=1}^{\infty} P(X=k)\,P(Y=k)$$

$$= \sum_{k=1}^{\infty} (q^{k-1}p)^2 = \sum_{t=0}^{\infty} q^{2t}p^2 \quad (t = k-1)$$

$$= p^2 \sum_{t=0}^{\infty} q^{2t} = \frac{p^2}{1-q^2}$$

Now, $(P_1 = P_2) \rightarrow P(X>Y) = P(Y>X)$

And $P(X>Y) + P(Y>X) + P(X=Y) = 1 \Rightarrow P(X>Y) = \frac{1-P(X=Y)}{2}$

$$= \frac{1 - \frac{p^2}{1-q^2}}{2} = \frac{1-p^2-q^2}{2(1-q^2)}$$

(5) Let each person has ~~an~~ unique id from $[1, n]$.

Let their ids are written on slip of paper.

let I_j be indicator r.v. for jth person draws jth slip.

From symmetry $P(I_j) = \frac{1}{n} \quad \forall\ 1 \leq j \leq n$

From linearity, $EX = EI_1 + EI_2 \cdots EI_n$

Where X is r.v. for total number of matches.

$$\Rightarrow \quad EX = \sum_{j=1}^{n} EI_j = n\, EI_1 = n \times \frac{1}{n} = 1$$

(6) Let after n tosses we get a string 's' of H's and T's of length n.

Let I_j be indicator r.v. for a HTH pattern in s starting at index j of s.

So, $P(I_j) = \frac{1}{8}$ (because there should H at index j, T at $j+1$ and H at $j+2$)

Let X be r.v. for ~~total~~ HTH patterns in s.

$$X = I_1 + I_2 + \cdots I_{n-2}$$

From linearity, $EX = EI_1 + EI_2 \cdots EI_{n-2}$

From symmetry, $EI_1 = EI_2 = \cdots EI_{n-2} = \frac{1}{8}$

$$\Rightarrow \quad EX = (n-2)\, EI_1 = \frac{n-2}{8}$$

7. (a) Let I_j be indicator r.v. for correct prediction of jth card.

A jth card no matter what prediction we make, Probability of it being correct is $\frac{1}{52}$

From linearity, symmetry and fundamental bridge,

$$EX = 52 \times EI_1 = 52 \times \frac{1}{52} = 1$$

(b) Lets say we are predicting jth card. Each time we predict correctly, sample space reduces by 1. let's say we predicted i cards correctly including jth card.

Let I_j be indicator r.v. for correct prediction of jth card.

$$\Rightarrow P(I_j) = \sum_{i=1}^{j} \left(\frac{1}{52} \cdot \frac{1}{51} \cdot \frac{1}{50} \cdots \frac{1}{52-i+1} \right) \binom{j-1}{i-1}$$

$$= \sum_{i=1}^{j} \frac{(52-i)!}{52!} \binom{j-1}{i-1}$$

Because there are (i-1) correctly predicted positions ~~expart~~ apart from jth position which can be chosen $\binom{j-1}{i-1}$ ways.

From linearity,

$$EX = \sum_{j=1}^{52} \sum_{i=1}^{j} \binom{j-1}{i-1} \frac{(52-i)!}{52!}$$

$$= \sum_{i=1}^{52} \sum_{j=i}^{52} \binom{j-1}{i-1} \frac{(52-i)!}{52!}$$

$$= \sum_{i=1}^{52} \frac{(52-i)!}{52!} \sum_{j=i}^{52} \binom{j-1}{i-1}$$

$$= \sum_{i=1}^{52} \frac{(52-i)!}{52!} \binom{52}{i} \qquad \left(\sum_{a=b}^{c} \binom{a}{b} = \binom{c+1}{b+1} \right)$$

$$= \sum_{i=1}^{52} \frac{1}{i!}$$

$$= \sum_{i=1}^{52} \frac{1^i}{i!} \quad , \text{ Let } S = \sum_{i=1}^{52} \frac{1^i}{i!}$$

Now, $\sum_{i=0}^{\infty} \frac{1^i}{i!} = e$

$$\Rightarrow 1 + \sum_{i=1}^{\infty} \frac{1^i}{i!} = e$$

$$\Rightarrow 1 + \sum_{i=1}^{52} \frac{1^i}{i!} + \sum_{i=53}^{\infty} \frac{1^i}{i!} = e$$

$$\Rightarrow S \approx e-1 \quad \left(\sum_{i=53}^{\infty} \frac{1^i}{i!} \to 0 \right)$$

(C) In this case for every new card sample space reduces by one.

So, $P(I_1) = \frac{1}{52}$, $P(I_2) = \frac{1}{51}$, $P(I_3) = \frac{1}{50}$ ... $P(I_j) = \frac{1}{52-j+1}$

From linearity,

$$EX = \sum_{j=1}^{52} \frac{1}{52-j+1} = \frac{1}{52} + \frac{1}{51} + \cdots ,$$

$$\approx \log_e 52 \approx 3.95$$

8.(a) Let there are n people

x = # people having birthday same as mine.

probability of one ~~each~~ person $= \frac{1}{365}$

From symmetry and linearity,

$$Ex = \frac{n}{365} \Rightarrow \lambda = \frac{n}{365}$$

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

required probability $= 1 - P(X=0)$

$$= 1 - e^{-\lambda} = \frac{1}{2}$$

$$\Rightarrow \lambda = \ln(2)$$

$$\Rightarrow n = 365 \ln(2) \approx 253$$

(b) Let I_{ij} be indicator r.v. for person i and j ($i<j$) having same birthday and hour.

$$P(I_{ij}) = \frac{1}{365 \times 24}$$

$$\Rightarrow \lambda = Ex = \binom{n}{2} \frac{1}{365 \times 24}$$ (as there are $\binom{n}{2}$ possibilities of i and j)

$$\Rightarrow \binom{n}{2} \frac{1}{365 \times 24} = \ln(2)$$

$$\Rightarrow n \approx 111$$

(c) Amongst 23 people who have 50% chance of atleast one birthday match there are $\binom{23}{2}$ pairs possible. Hence sample space for birth hour match is of the order of $\binom{23}{2}$ (i.e. $O(n^2)$) and not 23 (i.e. not $O(n)$).

Hence intuitively 24·23 does not appear to be correct answer.

Now to obtain same probabilites, λ value of the two distribution should be same.

$$\Rightarrow \binom{n_1}{2}\frac{1}{365\times 24} = \binom{n_2}{2}\frac{1}{365}$$

where n_1 is people required for birthday-birthhour match and n_2 for birthday match in order to obtain same probability.

$$\Rightarrow \binom{n_1}{2}\Big/\binom{n_2}{2} = 24$$

$$\Rightarrow \frac{n_1(n_1-1)}{n_2(n_2-1)} = 24 \Rightarrow \left(\frac{n_1}{n_2}\right)^2 \simeq 24$$

$$\Rightarrow \left(\frac{n_1}{n_2}\right) \simeq \sqrt{24} \simeq 5$$

(d) Let I_{ijk} be indicator r.v. for same birthday of person i, j and k ($i<j<k$).

$P(I_{ijk}) = \frac{1}{365^2}$ and there are $\binom{100}{3}$ triplets possible from linerity and symmetry,

$$\lambda = EX = \binom{100}{3} \frac{1}{365^2} \approx 1.213$$

$$\Rightarrow \quad P(X=0) = e^{-\lambda}$$

required probability $= 1 - e^{-\lambda}$

$$= 1 - e^{-(1.213)}$$
$$\approx 0.703$$

Now let Z_j be indicator r.v. that there is atleast one triplet whose birthday is on j^{th} day.

$$\Rightarrow P(Z_j) = 1 - \left(P(X=0) + P(X=1) + P(X=2) \right)$$

$$= 1 - \left(\left(\frac{364}{365}\right)^{100} + \binom{100}{1} \left(\frac{364}{365}\right)^{99} \frac{1}{365} + \binom{100}{2} \left(\frac{364}{365}\right)^{98} \left(\frac{1}{365}\right)^{2} \right)$$

$$= 1 - 0.997 \approx 0.0028$$

From linearity,

$$\lambda = EZ = 365 \times P(Z_j) \approx 1.023$$

Now, required probability $= 1 - e^{-\lambda}$

$$= 1 - e^{-1.023}$$
$$\approx 1 - 0.3593$$
$$\approx 0.6406$$

Second approach seems more accurate.