1

a Let E donates the excepted the number, we need to consider the first 2 throws

For the first throw, if the result is tail, whose probability is 0.5, then we need E more times, so the excepted number is $0.5 * (E + 1)$

If the first result is head and the second is tail, whose probability is 0.25, then we need E times, so the excepted number is $0.25 * (E + 2)$

If the first and second results are both head, whose probability is 0.25, then we need 0 times, so the excepted number is 0.25*2

Therefore $E = 0.5 * (E + 1) + 0.25 * (E + 2) + 0.25 * 2 \rightarrow E = 6$

b Let E donates the excepted the number, we need to consider the first n throws

For the first throw, similarly the excepted number is $0.5 * (E + 1)$

For the first two throws, similarly the excepted number is $0.25 * (E + 2)$

For the first n throws, the excepted number is $0.5^n * (E + n)$ since the probability to get $n - 1$ heads and 1 tails is 0.5^n

Therefore $E = \sum_{i=1}^{n} 0.5^i(E + i) + 0.5^n n \rightarrow E = \frac{\sum_{i=1}^{n} 0.5^i i + 0.5^n n}{1 - \sum_{i=1}^{n} 0.5^i}$

$1 - \sum_{i=1}^{n} 0.5^i = 1 - \frac{0.5(1 - 0.5^n)}{1 - 0.5} = 0.5^n$

$\sum_{i=1}^{n} 0.5^i i = 2(1 - 0.5^n - n0.5^{n+1}) = 2 - 0.5^{n-1} - 0.5^n n$

Then $E = 2^{n+1} - 2$

c We can consider the coin as a new coin with 8 results, one of them is THH and one is TTH, and we use n throws to get our first THH with getting HHH before, $\Pr(t = k) = (\frac{7}{8})^{(k-1)} * (1 - \frac{6}{7}^{k-1}) \frac{1}{8} = \frac{7^{k-1} - 6^{k-1}}{8^k}$

So the result is $\sum_{k=1}^{\infty} \frac{7^{k-1} - 6^{k-1}}{8^k} = \frac{1}{7} \sum_{k=1}^{\infty} (\frac{7}{8})^k - \frac{1}{6} \sum_{k=1}^{\infty} (\frac{3}{4})^k = \frac{1}{2}$

d The probability is same as the last question, which is $\frac{1}{2}$

e Similarly to problem e, the number depends on the length of given sequence n, and the excepted number $E = 2^{n+1} - 2$

2

Assume the probability of getting head is p,

The the likelihood function is $L(n : p) = p^n (1 - p)^{(100-n)}$

To maximize the likelihood function $\frac{d \ln(L(n:p))}{dp} = 0 \rightarrow n\frac{1}{p} + (100 - n)\frac{1}{p-1} = 0 \rightarrow \frac{100p - n}{p(p-1)} = 0 \rightarrow p = \frac{n}{100}$

Since $n = 100$, then $p = 1$

The value p follows the Beta distribution where $\alpha = 101, \beta = 1$.

3

The inverse may be extremely large for calculation as well as storage. Since the inverse will also be multiplied by the input feature data for regression, then any subtle difference in the input feature data will cause a large difference for the result.

In this case, I think we can just discard the eigenvalue and corresponding eigenvector from the original matrix, and use the remaining eigenvalue and eigenvector to reconstruct a new matrix in lower dimension. We also need to discard the corresponding feature in the input feature data during regression.

Since the eigenvalue is very small, which means the corresponding eigenvector have little effect on our regression result, hence we can choose to omit it.

4

a They follow the order statistics

b Through python code simulation, $E[Y] = \frac{1}{3}$ and $E[X] = \frac{2}{3}$, $E[\rho] = 0.5$

5

Assume x^3 ends with 51, which is only related to the value of last two digits of x. Assume the last two digits of x is ab, then $(ab)^3$ must end with 51. First b^3 must end with 1, therefore, $b = 1$
Similarly, $(a1)^3$ end with 51, which means $(a*10+1)^3 \% 100 = 51 \rightarrow (1000a^3 + 300a^2 + 30a + 1)\%100 = 51$
$(30a+1)\%100 = 51 \rightarrow a = 5$
Therefore, x must end with 51, and the probability of x^3 end with 51 is 0.51

6

The problem is equivalent that one point is fixed and we need to move another point

a $L = \frac{1}{2\pi}(\int_0^\pi \theta d\theta + \int_\pi^{2\pi}(2\pi - \theta)d\theta) = \frac{1}{\pi}\int_0^\pi \theta d\theta = \frac{\pi}{2}$

b Assume we consider the start point as the north pole of our sphere, then the second point can be considered as on a unit circle with the start point.
   Since the second point have same probability to be located on each ring, therefore, the length is same from question a, which is $\frac{\pi}{2}$