NAME: Kevin Motia
DATE: 2024-01-22
Homework: Assignment 1
Data Science II

Problem 1

There is no feature heading, time stamps, or units, so I wouldn't even be able to interpret what the data is of in the slightest if only given the text file. It also would've been useful if there was some documentation about how the data was collected.

Problem 2

2.1 Calculating the first moment of a binomial distribution:

$$\langle k \rangle = \sum_{k=0}^{n} \binom{n}{k} (1-p)^k p^{n-k} k$$

$$= \sum_{k=0}^{n} \frac{n!k}{k!(n-k)!} (1-p)^k p^{n-k}$$

$$= \sum_{k=0}^{n} \frac{n!k}{k!(n-k)!} (1-p)(1-p)^{k-1} p^{n-k}$$

$$= \sum_{k=0}^{n} \frac{n(n-1)!k}{k(k-1)!(n-k)!} (1-p)(1-p)^{k-1} p^{n-k}$$

$$= np \sum_{k-1=0}^{n-1} \frac{(n-1)!}{(k-1)!(n-k)!} (1-p)^{k-1} p^{n-k}$$

Using binomial theorem: $(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} x^k a^{n-k}$

$$\langle k \rangle = n(1-p)((1-p)+p)^{n-1}$$

$$= n(1-p)(1)$$

$$\langle k \rangle = n(1-p)$$

2.2 Calculating the second moment of a binomial distribution:

$$\langle k^2 \rangle = \sum_{k=0}^{n} \binom{n}{k} (1-p)^k p^{n-k} k^2$$

$$= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} (1-p)^k p^{n-k} ((k^2-k)+k)$$

$$= \sum_{k=0}^{n} \frac{n!(k^2-k)}{(n-k)!k!} (1-p)^k p^{n-k} + \sum_{k=0}^{n} \frac{n!k}{(n-k)!k!} (1-p)^k p^{n-k}$$

$$= n(n-1)(1-p)^2 \sum_{k-2=0}^{n-2} \frac{(n-2)!}{(k-2)!(n-k)!}(1-p)^{k-2} + n(1-p) \sum_{k-1=0}^{n-1} \frac{(n-1)!}{(k-1)!(n-k)!}(1-p)^{k-1}p^{n-k}$$

Using binomial theorem: $(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} x^k a^{n-k}$

$$< y^2 >= n(n-1)(1-p)^2((1-p)+(p))^{n-2} + n(1-p)((1-p)+p)^{n-1}$$

$$= n(n-1)(1-p)^2(1) + n(1-p)(1)$$

$$\langle k^2 \rangle = n(n-1)(1-p)^2 + n(1-p)$$

2.3 Determining $\langle y \rangle$:

$$\langle y \rangle = \sum_{y=0}^{\infty} y \sum_{x=y}^{\infty} P(x) \binom{x}{y} (1-p)^y p^{x-y}$$

$$= \sum_{x=0}^{\infty} \sum_{y=0}^{x} y P(x) \binom{x}{y} (1-p)^y p^{x-y}$$

First moment of binomial distribution:

$$\sum_{y=0}^{x} y \binom{x}{y} (1-p)^y p^{x-y} = x(1-p)$$

$$\therefore \langle y \rangle = \sum_{x=0}^{\infty} P(x) x (1-p)$$

$$\boxed{\langle y \rangle = \langle x \rangle (1-p)}$$

2.4 Determining $\langle y^2 \rangle$:

$$\langle y \rangle = \sum_{y=0}^{\infty} y^2 \sum_{x=y}^{\infty} P(x) \binom{x}{y} (1-p)^y p^{x-y}$$

$$= \sum_{x=0}^{\infty} \sum_{y=0}^{x} y^2 P(x) \binom{x}{y} (1-p)^y p^{x-y}$$

Second moment of binomial distribution:

$$\sum_{y=0}^{x} y^2 \binom{x}{y} (1-p)^y p^{x-y} = x(x-1)(1-p)^2 + x(1-p)$$

$$\therefore \langle y^2 \rangle = \sum_{x=0}^{\infty} P(x)(x(x-1)(1-p)^2 + x(1-p))$$

$$= \sum_{x=0}^{\infty}[P(x)x(x-1)(1-p)^2 + P(x)x(1-p)]$$

$$= \sum_{x=0}^{\infty}[P(x)(x^2-x)(1-p)^2 + P(x)x(1-p)]$$

$$= (\sum_{x=0}^{\infty} P(x)x^2 - \sum_{x=0}^{\infty} P(x)x)(1-p)^2 + \sum_{x=0}^{\infty} P(x)x(1-p)$$

$$\boxed{\langle y^2\rangle = (\langle x^2\rangle - \langle x\rangle)(1-p)^2 + \langle x\rangle(1-p)}$$

Problem 3

Expectation of a geometric distribution:

$$\langle x\rangle = \frac{1}{p}$$

Given that $p = \frac{1}{20}$

$$\langle x\rangle = 20$$

$\boxed{\text{We expect to ask 20 questions.}}$