Question 2, Assignment 4: CS 754, Spring 2024-25

Amitesh Shekhar
IIT Bombay
22b0014@iitb.ac.in

Anupam Rawat
IIT Bombay
22b3982@iitb.ac.in

Toshan Achintya Golla
IIT Bombay
22b2234@iitb.ac.in

April 04, 2025

1. Consider that there are n coupons, each of a different colour. Suppose we sample coupons uniformly at random and with replacement. Then answer the following questions from first principles. Do not merely quote results pertaining to any distribution other than Bernoulli. [3+3+3+3+(4+4 = 8)+2.5+2.5 = 25 points]

   (a) What is the probability q_j of getting a new coupon on the jth trial, assuming that new, unique coupons were obtained in all the previous trials? What is q_1?

   (b) If you were to toss a coin independently many times, what is the probability that the first head appears on the kth trial? Assume that the probability of getting a heads on any trial is q.

   (c) Let Y be the random variable which denotes the trial number on which the first head appears. Derive a formula for $E(Y)$ in terms of q.

   (d) Derive a formula for the variance of Y.

   (e) Let Z_n be a random variable denoting the number of trials by which each of the n different coupons were selected at least once. Applying result in previous parts, what is the expected value of Z_n? Derive an upper bound on the variance of Z_n. (You will need to use the following results: $\sum_{i=1}^{n} 1/i \approx \log n + \gamma + O(1/n)$ where $\gamma \approx 0.5772$ is a constant, and $\sum_{i=1}^{n} 1/i^2 < \sum_{i=1}^{\infty} 1/i^2 < \pi^2/6$).

   (f) Given the previous results, use Markov's inequality to upper bound $P(Z_n \geq t)$ for some value t.

   (g) Given the previous results, use Chebyshev's inequality to upper bound $P(Z_n \geq t)$ for some value t.

Soln:

(a) Since the coupons are sampled uniformly at random with replacement, each coupon has an equal probability of getting selected, i.e. $\frac{1}{n}$. Now, as per the given statement, at the time step 'j', all of the previously drawn coupons were distinct. This means, $(j-1)$ distinct coupons have been drawn till now. Thus, we have $n-(j-1)$ choices for the next coupon of the total n, to satisfy the criteria. Thus the probability that we get a new coupon on jth trial, assuming that all the previous coupons were unique as well is:

$$\boxed{q_j = \frac{n-j+1}{n}}$$

The special case, when $j = 1$, i.e. getting a unique coupon on the first trial is:

$$\boxed{q_1 = \frac{n-1+1}{n} = 1}$$

Since, no coupons have been drawn before, the coupon drawn for $j = 1$ will always be unique, and therefore its probability is 1.

(b) The probability of obtaining a head is q, thus the probability of obtaining a tail is $(1-q)$.
Since, the first head is obtained at kth trial, (k-1) trials must yield a tail.

$$\text{Pr(head for the first time in kth trial)} = \text{Pr(tails consecutively for k-1 trials)} \cdot \text{Pr(head at the kth trial)}$$

$$\boxed{\text{Pr(head for the first time in kth trial)} = (1-q)^{(k-1)} \cdot q}$$

(c) Y is the random variable representing the trial number associated with the appearance of the first head. From previous part:

$$\Pr(\text{Y = k}) = (1-q)^{k-1} \cdot q \text{ for k = 1, 2, 3, 4...}$$

We need to derive a formula for the Expected value of Y:

$$\mathbb{E}[Y] = \sum_{k=1}^{\infty} k \cdot \Pr(\text{Y = k}) = \sum_{k=1}^{\infty} k \cdot (1-q)^{k-1} \cdot q$$

$$\mathbb{E}[Y] = q \sum_{k=1}^{\infty} k \cdot (1-q)^{k-1}$$

See **Appendix A**, for the derivation of the identity:

$$\sum_{k=1}^{\infty} kx^k = \frac{1}{(1-x)^2} \text{ for } |x| < 1$$

Using the above result in the computation of Expected value (both q and 1-q are less than 1):

$$\mathbb{E}[Y] = q \sum_{k=1}^{\infty} k \cdot (1-q)^{k-1} = q \cdot \frac{1}{(1-(1-q))^2} = q \cdot \frac{1}{q^2} = \frac{1}{q}$$

$$\boxed{\mathbb{E}[Y] = \frac{1}{q}}$$

Thus, the expected number of trial on which the first head appears is $\frac{1}{q}$.
This makes sense, for the case where q = 0, in that case, (1-q) is not less than 1, and the sum never converges, meaning head never appears. And the case where q = 1, head appears on the first trial.

(d) The variance of a random variable Y is given by:

$$Var(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2$$

The first term can be computed in the following manner:

$$\mathbb{E}[Y^2] = \sum_{k=1}^{\infty} k^2 \cdot \Pr(\text{Y = k}) = \sum_{k=1}^{\infty} k^2 \cdot (1-q)^{k-1} \cdot q$$

Using the result derived in **Appendix B**, for $|x| < 1$. Substituting the result in the above computation:

$$\sum_{k=1}^{\infty} k^2 \cdot x^{k-1} = \frac{2x}{(1-x)^3} + \frac{1}{(1-x)^2}$$

$$\mathbb{E}[Y^2] = q \left(\sum_{k=1}^{\infty} k^2 \cdot (1-q)^{k-1} \right) = q \left(\frac{2(1-q)}{q^3} + \frac{1}{q^2} \right)$$

$$\mathbb{E}[Y^2] = \frac{2(1-q)}{q^2} + \frac{1}{q}$$

The variance can be computed by:

$$Var(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2$$

$$Var(Y) = \left(\frac{2-2q}{q^2} + \frac{1}{q} \right) - \left(\frac{1}{q} \right)^2 = \left(\frac{1-2q}{q^2} + \frac{1}{q} \right) = \frac{1-2q+q}{q^2}$$

The Variance is given by:

$$\boxed{Var(Y) = \frac{1-q}{q}}$$

(e) Z_n is the random variable denoting the number of trials to select n different coupons were selected atleast once. Let Y_j be the number of trials to get the j^th new coupon, given that $(j-1)$ coupons have already been collected. Then, Z_n is defined as:

$$Z_n = \sum_{j=1}^{n} Y_j$$

Using results derived from the previous parts
Each Y_j is a geometric random variable (from Bernoulli trials) with success probability:

$$q_j = \frac{n-j+1}{n}$$

Thus, expected value of Z_n is given by:

$$\mathbb{E}[Z_n] = \sum_{j=1}^{n} \mathbb{E}[Y_j] = \sum_{j=1}^{n} \frac{1}{q_j}$$

$$\mathbb{E}[Z_n] = \sum_{j=1}^{n} \frac{n}{n-j+1}$$

Substituting $k = n - j + 1$, the range transforms to $k = 1$ to $k = n$:

$$\mathbb{E}[Z_n] = \sum_{k=1}^{n} \frac{n}{k} = n \sum_{k=1}^{n} \frac{1}{k} = n \cdot H(n) \text{ , where H(n) } = \sum_{k=1}^{n} \frac{1}{k}$$

Substituting approximate value of H(n) from the given result:

$$\mathbb{E}[Z_n] \approx n\,(logn + \gamma + O(1/n))$$

$$\boxed{\mathbb{E}[Z_n] \approx n \cdot log(n) + n\gamma + O(1)}$$

The Variance of Z_n is given by:

$$Var(Z_n) = Var(\sum_{j=1}^{n} Y_j)$$

But since, the Y_j's are independent:

$$Var(Z_n) = \sum_{j=1}^{n} Var(Y_j) = \sum_{j=1}^{n} \left(\frac{1-q_j}{q_j^2} \right)$$

$$Var(Z_n) = \sum_{j=1}^{n} \left(\frac{1 - \frac{n-j+1}{n}}{(\frac{n-j+1}{n})^2} \right) = \sum_{j=1}^{n} \left(\frac{\frac{j-1}{n}}{(\frac{n-j+1}{n})^2} \right) = \sum_{j=1}^{n} \left(\frac{(j-1)n}{(n-j+1)^2} \right)$$

$$Var(Z_n) = n \sum_{j=1}^{n} \frac{j-1}{(n-j+1)^2}$$

Substituting $k = n - j + 1$, the range changes from $k = 1$ to $k = n$:

$$Var(Z_n) = n \sum_{k=1}^{n} \frac{n-k}{k^2} = n^2 \sum_{k=1}^{n} \frac{1}{k^2} - n \sum_{k=1}^{n} \frac{1}{k}$$

Substituting in the approximations, provided, we get the upper bound as:

$$\boxed{Var(Z_n) < n^2 \frac{\pi^2}{6} - n(log(n) + \gamma) + O(1)}$$

(f) From the Markov Inequality, we know that:

$$P(Z_n \geq t) \leq \frac{\mathbb{E}[Z_n]}{t}$$

Substituting the value of $\mathbb{E}[Z_n]$ from previous parts, we get:

$$\boxed{P(Z_n \geq t) \leq \frac{n(\log n + \gamma) + O(1)}{t}}$$

(g) Chebyshev's Inequality states that for any random variable X with finite mean μ and variance σ^2 and for any $k > 0$:

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

$$P(Z_n \geq t) = P(Z_n - \mu \geq t - \mu) \leq P(|Z_n - \mu| \geq t - \mu)$$

$$P(Z_n \leq t) \leq P(|Z_n - \mu| \geq t - \mu) \leq \frac{Var(Z_n)}{(t - \mu)^2}$$

Substituting the approximate value of Variance and Expected values from the previous parts, we get

$$\boxed{P(Z_n \leq t) \leq \frac{\pi^2 n^2}{6(t - nH(n))^2}}$$

A tighter and more accurate value would be:

$$P(Z_n \leq t) \leq \frac{n^2\pi^2 - 6n(\log(n) + \gamma) + 6O(1)}{6(t - n(\log(n) + \gamma) - O(1))^2}$$

Appendix

Appendix A

Proof of

$$\sum_{k=1}^{\infty} k \cdot x^{k-1} = \frac{1}{(1-x)^2}$$

We know that the infinite sum of GP when the common ratio is less than 1 is given by:

$$\sum_{k=1}^{\infty} x^k = \frac{1}{1-x}$$

Differentiating both sides w.r.t x, we get:

$$\sum_{k=1}^{\infty} \frac{d}{dx}\left(x^k\right) = \frac{d}{dx}\left(\frac{1}{1-x}\right)$$

$$\sum_{k=1}^{\infty} k \cdot x^{k-1} = \frac{(1-x)\frac{d}{dx}(1) - (1)\frac{d}{dx}(1-x)}{(1-x)^2}$$

$$\sum_{k=1}^{\infty} x^k = \frac{1}{1-x}$$

Hence Proved

Appendix B

Proof of

$$\sum_{k=1}^{\infty} k^2 \cdot x^{k-1} = \frac{2x}{(1-x)^3} + \frac{1}{(1-x)^2}, \text{ for } |x| < 1$$

From **Appendix A**,

$$\sum_{k=1}^{\infty} k \cdot x^{k-1} = \frac{1}{(1-x)^2}$$

Differentiating both sides once again w.r.t x:

$$\sum_{k=1}^{\infty} k(k-1) \cdot x^{k-2} = \frac{(1-x)^2\frac{d}{dx}(1) - 1\frac{d}{dx}(1-x)^2}{(1-x)^4}$$

$$\sum_{k=1}^{\infty} k(k-1) \cdot x^{k-2} = \frac{-(-2)(1-x)}{(1-x)^4} = \frac{2}{(1-x)^3}$$

Multiplying both sides by x, we obtain

$$\sum_{k=1}^{\infty} k(k-1) \cdot x^{k-1} = \frac{2x}{(1-x)^3}$$

$$\sum_{k=1}^{\infty} k^2 x^{k-1} - \sum_{k=1}^{\infty} k x^{k-1} = \frac{2x}{(1-x)^3}$$

Once again, using the result from **Appendix A**,

$$\sum_{k=1}^{\infty} k^2 x^{k-1} = \frac{2x}{(1-x)^3} + \frac{1}{(1-x)^2}$$

Hence Proved