$A_k = (1-p)(A_{k-1} + (1-p^{k-1})p \log_2 (1-p))$

$A_k = (1-p)^k p \cdot \log_2 (1-p)^k p$

2023Fall Probability & Statistics for EECS 2023/12/25

Homework 11

Professor: Ziyu Shao & Dingzhu Wen *Due: 2024/1/1 10:59pm*

1. Let X be a discrete r.v. whose distinct possible values are $x_0, x_1, \ldots$, and let $p_k = P(X = x_k)$. The entropy of X is $H(X) = \sum_{k=0}^{\infty} p_k \log_2(1/p_k)$.

   (a) Find $H(X)$ for $X \sim \text{Geom}(p)$.

   (b) Let X and Y be i.i.d. discrete r.v.s. Show that $P(X = Y) \geq 2^{-H(X)}$. **Hint**: Jensen's Inequality.

   $e^{-\lambda}$

2. Let $X \sim \text{Pois}(\lambda)$. The conditional distribution of X, given that $X \geq 1$, is called a truncated Poisson distribution.

   $E[X|X\geq 1] = \sum_{k=0}^{\infty} k$ —

   (a) Find $E(X|X \geq 1)$.

   (b) Find $\text{Var}(X|X \geq 1)$.

3. Let $X_1 \sim \mathit{Expo}(\lambda_1)$, $X_2 \sim \mathit{Expo}(\lambda_2)$ and $X_3 \sim \mathit{Expo}(\lambda_3)$ be independent.

   $= \frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!}$

   (a) Find $E(X_1|X_1 > 2023)$

   (b) Find $E(X_1 + X_2 + X_3|X_1 > 2023, X_2 > 2024, X_3 > 2025)$ in terms of $\lambda_1, \lambda_2, \lambda_3$.

4. Let X and Y be two continuous random variables with joint PDF

   $= \frac{\lambda e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$

$$f_{X,Y}(x, y) = \begin{cases} 6xy & \text{if } 0 \leq x \leq 1, 0 \leq y \leq \sqrt{x}, \\ 0 & \text{otherwise.} \end{cases}$$

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

   (a) Find the marginal distributions of X and Y. Are X and Y independent?

   (b) Find $E[X|Y = y]$ and $\text{Var}[X|Y = y]$ for $0 \leq y \leq 1$.

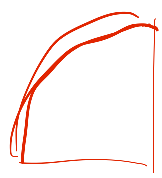
   (c) Find $E[X|Y]$ and $\text{Var}[X|Y]$.

5. Instead of predicting a single value for the parameter, we give an interval that is likely to contain the parameter: A $1 - \delta$ confidence interval for a parameter p is an interval $[\hat{p} - \epsilon, \hat{p} + \epsilon]$ such that $Pr\left(p \in [\hat{p} - \epsilon, \hat{p} + \epsilon]\right) \geq 1 - \delta$. Now we toss a coin with probability p landing heads and probability $1 - p$ landing tails. The parameter p is unknown and we need to estimate its value from experiment results. We toss such coin N times. Let $X_i = 1$ if the ith result is head, otherwise 0. We estimate p by using

$\frac{1}{N}$ Bin

$$\hat{p} = \frac{X_1 + \ldots + X_N}{N}.$$

$\frac{p(1-p)}{n}$



   Find the $1 - \delta$ confidence interval for p, then discuss the impacts of δ and N.

$1-\delta$ $P(|\hat{p} - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}$ δ $1-\delta$

$P(|\hat{p} - p| < \varepsilon) \geq 1 - \delta$

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

(a) Method 1: Adopt Chebyshev inequality to find the $1-\delta$ confidence interval for p, then discuss the impacts of δ and N.

(b) Method 2: Adopt Hoeffding bound to find the $1-\delta$ confidence interval for p, then discuss the impacts of δ and N.

(c) Discuss the pros and cons of the above two methods. $\geq 2e^{-2N\varepsilon^2}$

E

6. (**Optional Challenging Problem**) A coin with probability p of Heads is flipped repeatedly. For (a) and (b), suppose that p is a known constant, with $0 < p < 1$.

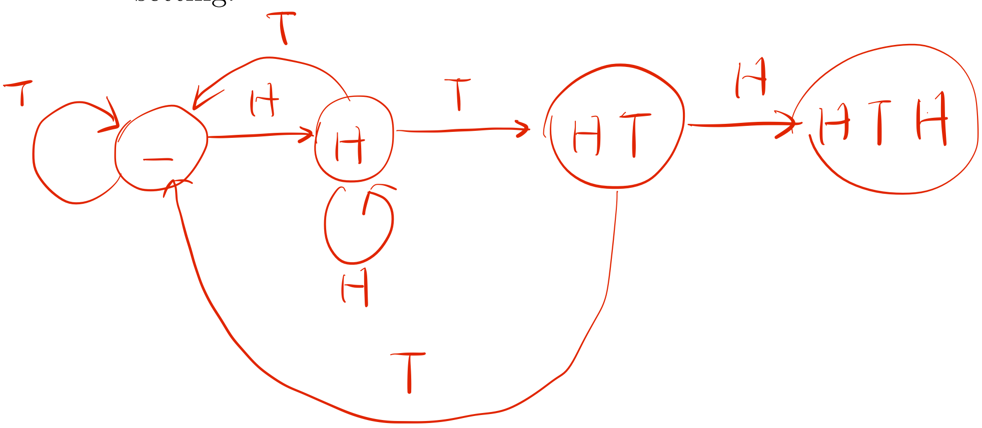
δ

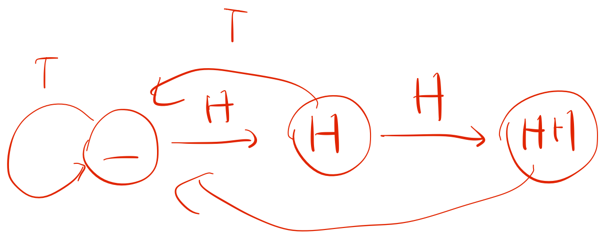
(a) What is the expected number of flips until the pattern HT is observed?

(b) What is the expected number of flips until the pattern HH is observed?

(c) What is the expected number of flips until the pattern HTH is observed?

(d) Now suppose that p is unknown, and that we use a Beta(a, b) prior to reflect our uncertainty about p (where a and b are known constants and are greater than 2). In terms of a and b, find the corresponding answers to (a), (b) and (c) in this setting.





$E(M) = E(2N+2) = 10$

$E(M|N) = 2N+2$

$E(M|N=k)$ $\quad E(M) = ?$

$= \frac{1}{2}(k+1) + \frac{1}{2}(k+1+E(M|N=k))$

4

6