CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair

HW 13

Due: Saturday, 4/20, 4:00 PM
Grace period until Saturday, 4/20, 6:00 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Uniform Uniform Computation

Note 21

Suppose $X \sim \text{Uniform}[0,1]$ and $Y \sim \text{Uniform}[0,X]$. That is, conditioned on $X = x$, Y has a $\text{Uniform}[0,x]$ distribution.

(a) What is $\mathbb{P}[Y > 1/2]$?

(b) Calculate $\text{Cov}(X,Y)$.

2 Moments of the Gaussian

Note 21

For a random variable X, the quantity $\mathbb{E}[X^k]$ for $k \in \mathbb{N}$ is called the *kth moment* of the distribution. In this problem, we will calculate the moments of a standard normal distribution.

(a) Prove the identity

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{tx^2}{2}\right) \mathrm{d}x = t^{-1/2}$$

for $t > 0$.

Hint: Consider a normal distribution with variance $\frac{1}{t}$ and mean 0.

(b) For the rest of the problem, X is a standard normal distribution (with mean 0 and variance 1). Use part (a) to compute $\mathbb{E}[X^{2k}]$ for $k \in \mathbb{N}$.

Hint: Try differentiating both sides with respect to t, k times. You may use the fact that we can differentiate under the integral without proof.

(c) Compute $\mathbb{E}[X^{2k+1}]$ for $k \in \mathbb{N}$.

1. (a) $cdf(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$ $pdf(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \le x \le 1 \\ 0 & x > 1 \end{cases}$

$cdf(y) = \begin{cases} 0 & y < 0 \\ y & 0 \le y < x \\ 0 & y \ge x \end{cases}$ $pdf(y) = \begin{cases} 0 & y < 0 \\ 1 & 0 \le y < x \\ 0 & y > x \end{cases}$

$P[Y > \frac{1}{2}]$ $\underline{?}$

$$P[Y > \tfrac{1}{2} \mid X = x] = \begin{cases} 0 & x < \frac{1}{2} \\ \dfrac{x - \frac{1}{2}}{x} & 1 > x \ge \frac{1}{2} \end{cases}$$

$$P[Y > \tfrac{1}{2}] = \int_{-\infty}^{+\infty} P[Y > \tfrac{1}{2} \mid X = x] f_X(x)\, dx$$

$$= \int_{\frac{1}{2}}^{1} \left(1 - \frac{1}{2x}\right) dx = x - \frac{1}{2}\ln x \Big|_{\frac{1}{2}}^{1} = \frac{1}{2}(1 - \ln 2)$$

(b) $Cov(X, Y) = E(XY) - E(X)E(Y)$

$$E(x) = \int_0^1 x\, dx = \frac{1}{2}x^2 \Big|_0^1 = \frac{1}{2}$$

$$E(Y) = \int_0^1 \int_0^x y\, dy\, dx = \int_0^1 \frac{1}{2}y^2 \Big|_0^x dx = \int_0^1 \frac{1}{2}x^2\, dx = \frac{1}{6}x^3 \Big|_0^1 = \frac{1}{6}$$

$$E(XY) = \int_0^1 x P[Y = y \mid X = x] f_X(x)\, dx$$

$$= \int_0^1 x \int_0^x y\,dy\,dx = \int_0^1 \frac{1}{2}x^3\,dx = \frac{1}{8}x^4\Big|_0^1 = \frac{1}{8}$$

$$Cov(X, Y) = \frac{1}{8} - \frac{1}{2} * \frac{1}{6} = \frac{1}{8} - \frac{1}{12} = \frac{1}{24}$$

2.(a) $$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}$$

$$\mu = 0 \qquad \sigma^2 = \frac{1}{t}$$

$$f(x) = \sqrt{\frac{t}{2\pi}}\, e^{-\frac{tx^2}{2}}$$

$$\int_{-\infty}^{+\infty} f(x) = 1 \Rightarrow \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{tx^2}{2}}\,dx = t^{-\frac{1}{2}}$$

(b) $$E(X^{2k}) = \int_{-\infty}^{+\infty} x^{2k} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$$

$$= \int_{-\infty}^{+\infty} \frac{1}{2k+1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,d\,x^{2k+1}$$

$$= \frac{1}{2k+1}\left[\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot x^{2k+1} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} x^{2k+1}\,d\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right.$$

???

(c) 0 density function is symmetric around 0.

(b) $\dfrac{d^k}{dt^k}\left[\dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \exp\left(-\dfrac{tx^2}{2}\right)dx\right] = \dfrac{d^k}{dt^k}\left(t^{-\frac{1}{2}}\right)$

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(-\frac{x^2}{2}\right)^k \exp\left(-\frac{tx^2}{2}\right)dx$$

LHS:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} (-1)^k * \frac{x^{2k}}{2^k} \exp\left(-\frac{tx^2}{2}\right)dx$$

RHS: $\left(-\frac{1}{2}\right) t^{-\frac{3}{2}}$

$$(-1)^k * \frac{1}{2^k} * 1 * 3 * 5 * \cdots (2k-1)\, t^{-\frac{2k+1}{2}}$$

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^{2k} \exp\left(-\frac{tx^2}{2}\right)dx = 1*3*5*(2k-1)\, t^{-\frac{2k+1}{2}}$$

$t=1$

$$E(X^{2k}) = \prod_{i=1}^{k}(2i-1)$$

3 Exponential Median

Note 21

(a) Prove that if $X_1, X_2, \ldots, X_n$ are mutually independent exponential random variables with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$, then $\min(X_1, X_2, \ldots, X_n)$ is exponentially distributed with parameter $\sum_{i=1}^{n} \lambda_i$.

Hint: Recall that the CDF of an exponential random variable with parameter λ is $1 - e^{-\lambda t}$.

(b) Given that the minimum of three i.i.d exponential variables with parameter λ is m, what is the probability that the difference between the median and the smallest is at least s? Note that the exponential random variables are mutually independent.

(c) What is the expected value of the median of three i.i.d. exponential variables with parameter λ?

Hint: Part (b) may be useful for this calculation.

4 Chebyshev's Inequality vs. Central Limit Theorem

Note 17
Note 21

Let n be a positive integer. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with the following distribution:

$$\mathbb{P}[X_i = -1] = \frac{1}{12}; \qquad \mathbb{P}[X_i = 1] = \frac{9}{12}; \qquad \mathbb{P}[X_i = 2] = \frac{2}{12}.$$

(a) Calculate the expectations and variances of X_1, $\sum_{i=1}^n X_i$, $\sum_{i=1}^n (X_i - \mathbb{E}[X_i])$, and

$$Z_n = \frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{n/2}}.$$

(b) Use Chebyshev's Inequality to find an upper bound b for $\mathbb{P}[|Z_n| \geq 2]$.

(c) Use b from the previous part to bound $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$.

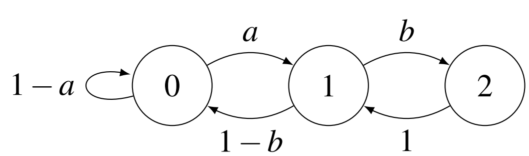
(d) As $n \to \infty$, what is the distribution of Z_n?

(e) We know that if $Z \sim \mathcal{N}(0,1)$, then $\mathbb{P}[|Z| \leq 2] = \Phi(2) - \Phi(-2) \approx 0.9545$. As $n \to \infty$, provide approximations for $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$.

5 Analyze a Markov Chain

Note 22

Consider a Markov chain with the state diagram shown below where $a, b \in (0,1)$.



3(a) $f(x) = \lambda e^{-\lambda x}$

$\overset{A}{\min}(X_1, X_2 \ldots X_n) = x$

$P(A \geq x) = P(X_1 \geq x \cap X_2 \geq x \cap x_3 \geq x \cap \cdots x_n \geq x)$

$= P(X_1 \geq x) * P(X_2 \geq x) \ldots P(X_n \geq x)$

$P(X_1 \geq x) = \int_x^{+\infty} \lambda e^{-\lambda x_1} = -e^{-\lambda x_1}\Big|_x^{+\infty} = e^{-\lambda_1 x}$

$P(A \geq x) = e^{-(\lambda_1 + \lambda_2 + \ldots \lambda_n) x} = \int_x^{+\infty} f(x)\, dx$

$\underline{P(A \leq x)} = 1 - e^{-(\lambda_1 + \lambda_2 + \ldots \lambda_n) x}$

$\underline{\text{cdf}}$ $\prime\prime$

$\text{pdf}(x) = (\lambda_1 + \lambda_2 + \ldots \lambda_n)\, e^{-(\lambda_1 + \lambda_2 + \lambda_n) x} \sim Exp(\sum_{i=1}^{n} \lambda_i)$

(b) $\underline{???}$ | X_1 is the minimum ,

$P(\min(X_2, X_3) - X_1 > s)$

$= P(\min(X_2, X_3) > m+s \mid X_1 = m)$ independ X_1, X_2, X_3

$= P(\min(X_2, X_3) > m+s) = P(X_2 > m+s) * P(X_3 > m+s)$

$= e^{-\lambda(m+s)} * e^{-\lambda(m+s)} = \underline{e^{-2\lambda(m+s)}}$

(c) $E(\min(X_2, X_3))$ (b) set $s=0$

$e^{-2\lambda m}$

???

$E(\min) = \frac{1}{3\lambda}$ from (a)

$E(\underbrace{\overset{s}{\text{med}-\min}}_{A}) = \frac{1}{2\lambda}$ → proof? → very complex

$E(\text{med}) = \frac{1}{3\lambda} + \frac{1}{2\lambda} = \frac{5}{6\lambda}$

$$f_A(s) = \frac{d(1-e^{-2\lambda(m+s)})}{ds} = 2\lambda e^{-2\lambda(m+s)}$$

$$E(A) = \int_0^{+\infty}\int_0^{+\infty} 2\lambda e^{-2\lambda(m+s)}\, ds\, dm$$

$$= \int_0^{+\infty} -e^{-2\lambda(m+s)}\Big|_0^{+\infty} dm = \int_0^{+\infty} e^{-2\lambda m}\, dm$$

$$= -\frac{1}{2\lambda}e^{-2\lambda m}\Big|_0^{+\infty} = \frac{1}{2\lambda}$$ good

4. (a) $E(X_1) = \frac{1}{12}\times -1 + \frac{9}{12}\times 1 + \frac{2}{12} * 2 = 1$

$$E(X_1^2) = \frac{1}{12}\times 1 + \frac{9}{12}\times 1 + \frac{2}{12}\times 4 = \frac{18}{12} = \frac{3}{2}$$

$$Var(X_1) = E(X_1^2) - (E(X_1))^2 = \frac{3}{2} - 1 = \frac{1}{2}$$

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = n$$

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = \frac{n}{2}$$

$$E\left(\sum_{i=1}^{n}(x_i - E(x_i))\right) = n - n = 0 \quad E(Z_n) = 0$$

$$Var\left(\sum_{i=1}^{n}[x_i - E(x_i)]\right) = \frac{n}{2} \quad Var(Z_n) = \frac{\frac{n}{2}}{\frac{n}{2}} = 1$$

(b) $P[|X-\mu| \geq c] \leq \frac{Var(X)}{c^2}$

$$P[|Z_n - 0| \geq 2] \leq \frac{Var(Z_n)}{4} = \underline{\frac{1}{4}}$$

$b = \frac{1}{4}$

$$x_i - E(x_i) = \begin{cases} -2 & \frac{1}{12} \\ 0 & \frac{9}{12} \\ 1 & \frac{2}{12} \end{cases}$$

(c) $\underline{\leq \frac{b}{2}}$ ⟷ upper boundary ↓ $\frac{1}{4}$ for both

(d) $\underline{N(0,1)}$

(f) $\frac{(1-0.9545)}{2} = \underline{0.02275}$

Here, we let $X(n)$ denote the state at time n.

(a) Show that this Markov chain is aperiodic.

(b) Calculate $\mathbb{P}[X(1)=1, X(2)=0, X(3)=0, X(4)=1 \mid X(0)=0]$.

(c) Calculate the invariant distribution.

5 : (a) There is a loop at 0, so it's aperiodic.

(b) $P[X(1)=1, X(2)=0, X(3)=0, X(4)=1 \mid X(0)=0]$

$= P(X(1)=1 \mid X(0)=0) * P(X(2)=0 \mid (X(1)=1, X(0)=0)) \cdot \ldots$

$= a * (1-b)(1-a)\, a = a^2(1-a)(1-b)$

(c) $(\pi_0, \pi_1, \pi_2) = (\pi_0, \pi_1, \pi_2) \begin{bmatrix} 1-a & a & 0 \\ 1-b & 0 & b \\ 0 & 1 & 0 \end{bmatrix}$

$\pi_0 + \pi_1 + \pi_2 = 1$

$\pi_0 = \pi_0(1-a) + \pi_1(1-b)$

$\pi_1 = a\pi_0 + \pi_2$

$\pi_2 = b\pi_1$

$(1-b)\pi_1 = a\pi_0$

$\pi_2 = b\pi_1$

$\frac{1-b}{a}\pi_1 + \pi_1 + b\pi_1 = 1$

$\pi_1\left(\frac{1-b+a+ab}{a}\right) = 1$

$\left(\frac{1-b}{a} + \frac{a}{a} + \frac{ab}{a}\right)$

$\pi_1 = \frac{a}{ab+a-b+1}$ $\pi_2 = \frac{ab}{ab+a-b+1}$

$\pi_0 = \frac{1-b}{ab+a-b+1}$