15-451: Algorithms Sept 24, 2019

Lecture Notes: Probability Review

Lecturer: Gary Miller *Scribe:*

[1]

1 Depth-first search basics

2 The Exponential Distribution

Definition 2.1. Let Ω be a sample space, a random variable is a mapping $X : \Omega \to \mathbb{R}$.

Definition 2.2. The probability density distribution (PDF) of an exponential random variable X_β is

$$\Pr[X_\beta = \mu] = \begin{cases} \beta e^{-\beta\mu}, & \mu \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Definition 2.3. The culmulutive distribution function (CDF) of X_β is

$$F_\beta(y) \equiv \Pr[X_\beta \leq y]$$
$$F_\beta(y) = \int_0^y \beta e^{-\beta x} dx = [-e^{-\beta x}]_0^y = 1 - e^{-\beta y}$$

Definition 2.4. The expected value of a random variable X is

$$\mathbb{E}_x[X] = \int_{-\infty}^{\infty} y \Pr[X = y] dy$$

Remark 2.5. There are two ways to calculate $\mathbb{E}[X_\beta]$ for a exponential random variable X_β

1. By defintion, using integration by parts,

$$\mathbb{E}[X_\beta] = \int_0^\infty y\beta e^{-\beta y} dy = 1/\beta$$

2.

$$\mathbb{E}[X_\beta] = \int_0^\infty \Pr[X_\beta \geq y] dy = \int_0^\infty e^{-\beta y} = [-\frac{1}{\beta} e^{-\beta y}]_0^\infty = \frac{1}{\beta}$$

Proposition 2.6 (Memoryless Property)**.** *Given exponential random variable X_β,*

$$\Pr[X_\beta > m + n | X_\beta > n] = \frac{e^{-\beta(m+n)}}{e^{-\beta n}} = e^{-\beta m}$$

[1]Originally 15-750 notes by Andre Wei

3 Order Statistics

Definition 3.1. $X_1, X_2, \dots X_n$ are n i.i.d random variables. The i-th order statistic is

$$X_{(i)} = \text{SELECT}_k(X_1, \dots X_k)$$

i.e.

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}.$$

Theorem 3.2. *Suppose $X_1, X_2, \dots, X_n$ are i.i.d such that*

$$f(u) = \Pr[X_i = u]$$

and

$$Fu = \Pr[0 \leq X_i \leq u].$$

Then

$$\Pr[X_{(1)} = u] = n(1 - F(u))^{n-1} f(u)$$

Corollary 3.3. *If $X_1, X_2, \dots X_n$ are i.i.d exponentials,*

$$\Pr[X_{(1)} = u] = n(e^{-\beta u})^{n-1} \beta e^{-\beta u} = n\beta e^{-n\beta u}$$

So $X_{(1)} \sim Exp(n\beta)$. Therefore

$$\mathbb{E}(X_{(1)}) = \frac{1}{n\beta}.$$

Claim 3.4 (Expectation of $X_{(n)}$)**.**

$$X_{(n)} \approx \frac{\log n}{\beta}$$

Proof. Let $S_i = X_{(i+1)} - X_{(i)}$, for $i \geq 0$.
We will need the following sub-claim:

Claim 3.5.

$$S_i \sim Exp((n-i)\beta)$$

We will prove this claim using the memoryless property. We think of each X_i as a time, say, the time that the ith light bulb burnt out. Thus at time $X_{(i)}$ i of the bulbs have burnt out and $n - i$ still lit. Assume that the burnt-out ones are $X_1, \dots X_i$, thus $X_j > X_{(i)}$ for $i < j \leq n$. Thus $S_i \sim X_{(1)}$ but for $n - i$ random variables. □

Thus,

$$\mathbb{E}(S_i) = \frac{1}{(n-i)\beta}$$

Therefore,

$$\mathbb{E}(X_{(n)}) = \sum_{i=0}^{n-1} \mathbb{E}[S_i] = \frac{1}{\beta}(1 + \frac{1}{2} + \dots \frac{1}{n}) = \frac{\ln n}{\beta}$$

□

Proposition 3.6 (Concentration for $X_{(n)}$)**.**

$$\Pr[X_i \geq \frac{c \ln n}{\beta}] = e^{-c \ln n} = n^{-c}$$

By union bound we get,

$$\Pr[X_i \geq \frac{c \ln n}{\beta}] \leq n \cdot n^{-c} = \frac{1}{n^{c-1}}$$

Thus,

$$\Pr[X_i \geq \frac{2 \ln n}{\beta}] \leq \frac{1}{n}$$

4 Generating Distribution of Random Variables

Problem: Given $f : \mathbb{R} \to \mathbb{R}^+$, where

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

Want to find random variable X_f whose PDF is f.

Remark 4.1. It is not clear that the random variable exists. But we can ask if we have one, can we generate more.

Definition 4.2. Let f, g be PDF's with random variable X_f, X_g, we say $f \leq g$ if there exists a deterministic process D such that $X_f = D(X_g)$.

Example 4.3. Let U be uniform random variable with PDF u, i.e.

$$u(x) = \begin{cases} 1, & \text{if } x \in [0,1], \\ 0, & \text{otherwise.} \end{cases}$$

Let U_2 be uniform random variable on $[0, 2]$,with PDF u_2, then

$$U_2 = 2U \implies u_2 \leq u$$

4.1 Generating Exponential Distribution from Uniform Distribution

The PDF of an exponential random variable X is

$$f(X) = \beta e^{-\beta X} \quad \text{for } 0 < \beta, X \geq 0$$

and

$$F(X) = \int_0^{\infty} f(X) dX = 1 - e^{-\beta X}$$

Thus $F : [0, \infty] \to [0, 1]$ is one-to-one and onto. We get that $F(X_f)$ is uniform on $[0, 1]$.
Therefore, $u \leq f$, But we want $f \leq u$.

Find F^{-1}, i.e. solve for X in $Y = F(X) = 1 - e^{-\beta X}$

$$\begin{aligned}
& Y = 1 - e^{-\beta X} \\
\iff\quad & e^{-\beta X} = 1 - Y \\
\iff\quad & -\beta X = \ln(1 - Y) \\
\iff\quad & X = -\frac{1}{\beta}\ln(1 - Y) \\
\iff\quad & X = -\frac{1}{\beta}\ln Y \quad \text{since } 1 - Y \text{ is uniform on } [0,\, 1]
\end{aligned}$$

Thus $X_f = \frac{1}{\beta}\ln(X_u)$. Thus $f \leq u$.

4.2 Generating Normal Distribution from Uniform Distribution

The PDF of a general normal random variable X is

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{X^2}{2\sigma^2}}$$

Taking $\sigma = 1$, we get Gauss' unit normal:

$$f(X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{X^2}{2}}$$

But it is hard to compute the CDF of X

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

Theorem 4.4. *F(X) is not an elementary function.*

Remark 4.5. It is OK to compute if $f(x) = xe^{-\frac{x^2}{2}}$, as

$$\frac{d}{dx}(-e^{-\frac{x^2}{2}}) = xe^{-\frac{x^2}{2}}$$

We consider 2D-normal.

$$\begin{aligned}
\text{Let } f(x, y) &= \frac{1}{2\pi} e^{-\frac{x^2}{2}} e^{-\frac{y^2}{2}} \\
&= \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}
\end{aligned}$$

In polar,

$$f(r, \theta) = \frac{1}{2\pi} e^{-\frac{r^2}{2}}$$

Now we can find the cumulative with respect to a disk of radiu r:

$$D(R) = \int_0^R \frac{2\pi r}{2\pi} e^{-\frac{r^2}{2}}\, dr = -e^{-\frac{r^2}{2}}]_0^R = 1 - e^{-\frac{R^2}{2}}$$

Again we compute F^{-1},

$$\begin{aligned}
\text{Let } y &= 1 - e^{-\frac{R^2}{2}} \\
\implies e^{-\frac{R^2}{2}} &= 1 - y \\
\implies -\frac{R^2}{2} &= \ln(1-y) \\
\implies R&\sqrt{-2\ln(1-y)}
\end{aligned}$$

Therefore given two uniform random variables u, v, we can generate a unit normal random variable using the following algorithm.

Alg: u, v uniform on $[0, 1]$.
  $r = \sqrt{-2\ln u}$
  $\theta = 2\pi v$
  In polar, **return** (r, θ)
  (or **return** $(x = r\cos\theta, y = r\sin\theta)$)

4.3 The Box-Muller Algorithm

Alg BM(u, v): u, v uniform on $[0, 1]$.
  1) Set $u = 2u - 1$, $v = 2v - 1$, (uniform on $[-1, 1]$)
  2) **do** $w = u^2 + v^2$ **until** $w \leq 1$
  3) Set $A = \sqrt{\frac{-2\ln w}{w}}$
  4) **return** $(T_1 = Au, T_2 = Av)$

Claim 4.6. *The Box-Muller Algorithm generates 2D unit Gaussian.*

Proof. After step 2), write u, v as

$$\begin{aligned}
V_1 &= R\cos\theta \\
V_2 &= R\sin\theta \\
S &= R^2
\end{aligned}$$

After step 4), we get the coordinate (x_1, x_2) where

$$x_1 = \sqrt{\frac{-2\ln S}{S}} V_1 = \sqrt{\frac{-2\ln S}{S}} R\cos\theta = \sqrt{-2\ln S}\cos\theta$$

Similarly,

$$X_2 = \sqrt{-2\ln S}\sin\theta$$

In polar form, we have (R', θ'), where $R' = \sqrt{-2\ln S}$, $\theta' \in [0, 2\pi]$.
Compute CDF of R',

$$\begin{aligned}
CDF(R') &= \Pr[R' \leq r] \\
&= \Pr[\sqrt{-2\ln S} \leq r] \\
&= \Pr[-2\ln S \leq r^2] \\
&= \Pr[S \geq e^{r^2/2}](*)
\end{aligned}$$

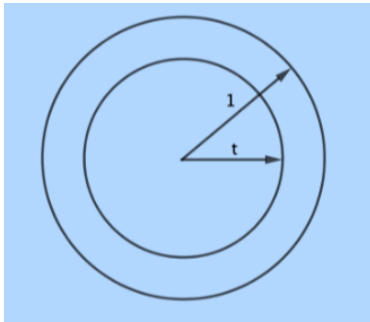

Figure 1: Visualization of $r \geq t$

Note suppose u, v is uniform over the unit disk, then in the figure below,

$$Pr[(u, v) \in \text{annulus}] = 1 - t^2$$

Consider random variable $S = R^2 = u^2 + v^2$,

$$\Pr[S \geq t] = \Pr[R^2 \geq t] = \Pr[R \geq \sqrt{t}] = 1 - t$$

Therefore,

$$\Pr[S \geq e^{\frac{r^2}{2}}] = 1 - e^{\frac{r^2}{2}}$$

So S is Gaussian. This completes our proof. $\square$