Probability 101

15-750
2/15/19

Random Variable: Samples $\to \mathbb{R}$

The Exponential Distribution

Prob Density Fcn

Def The Exponential RV X_β.

$$\text{Prob}[X_\beta = \mu] = \begin{cases} \beta e^{-\beta\mu} & \mu \geq 0 \\ 0 & \text{o.w} \end{cases}$$

Cumulative Dist

Def $F_\beta(y) \equiv \text{Prob}[X_\beta \leq y]$

$$F_\beta(y) = \int_0^y \beta e^{-\beta x}\,dx \equiv \left[-e^{-\beta x}\right]_0^y = 1 - e^{-\beta y}$$

Expected Value

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

Two ways to compute $\mathbb{E}[X_\beta]$.

1) **Def** $\mathbb{E}[X_\beta] = \int_0^{y} y\beta e^{-\beta y}\, dX = 1/\beta$

Using integration by parts.

2) $\mathbb{E}[X_\beta] = \int_0^{\infty} \text{Prob}[X_\beta \geq y]\, dy$

Now $\text{Prob}[X_\beta \geq y] = e^{-\beta y}$

thus $\mathbb{E}[X_\beta] = \int_0^{\infty} e^{-\beta y} = -\frac{1}{\beta} e^{-\beta y}\Big]_0^{\infty} = 1/\beta$

Memoryless Property

$$\text{Prob}[X_\beta > m+n \mid X_\beta > n] = \frac{e^{-\beta(n+m)}}{e^{-\beta n}} = e^{-\beta m}$$

Order Statistics

Suppose $X_1, \cdots, X_n$ random variables

Def $X_{(k)} \equiv \text{Select}_k(X_1, \cdots, X_n)$

ie $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$

Let $X_1, \ldots, X_n$ be iid exponentials.

Suppose $X_1, \dots, X_n$ are iid st $f(u) = \text{Prob}[X_i = u]$

and $F(u) = \text{Prob}[0 \leq X_i \leq u]$

<u>Thus</u> $X_{(1)} = n(1 - F(u))^{n-1} f(u)$

If $X_1, \dots, X_n$ are iid exponentials.

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

$$\equiv \text{Exp}(n\beta) \quad \text{Thus } \mathbb{E}\text{xp}[X_{(1)}] = \frac{1}{n\beta}$$

Let $S_i = X_{(i+1)} - X_{(i)}$ for $i \geq 0$

By Memoryless property $S_i = \text{Exp}((n-i)\beta)$

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

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

$$\approx \frac{\ln n}{\beta}$$

Concentration for $X_{(n)}$?

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

by union bound we get.

$$\Pr\left[X_{(n)} \geq \frac{c \ln n}{\beta}\right] \leq n \cdot n^{-c} = \frac{1}{n^{(c-1)}}$$

thus

$$\Pr\left[X_{(n)} \geq \frac{2 \ln n}{\beta}\right] \leq \frac{1}{n}$$

Generating Dist or Random Variables

Let X_f be the random variable of PDF $f: \mathbb{R} \to \mathbb{R}^+$ where $\int_{-\infty}^{\infty} f(x)dx = 1$

<u>Note</u> Not clear that RV exist.

But we can ask if we have one can we generate more.

<u>Def:</u> $f, g \in$ PDF's with RV Variables X_f & X_g

we say $f \leq g$ if $\exists$ det process D s.t. $X_f = D(X_g)$

Let $\mathcal{U}$ be uniform RV with PDF u.

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

Let $\mathcal{U}_2$ be uniform $[0,2]$ then

$X_{u_2} = 2X_u$ thus $\mathcal{U}_2 \leq \mathcal{U}$

Generating Exponential Dist from Uniform

PDF $\equiv f(x) = \beta e^{-\beta x}$ for $0<\beta$ & $x \geq 0$

$$F(x) = \int_0^x f(x)\,dx = 1 - e^{-\beta x}$$

Thus $F: [0,\infty] \longrightarrow [0,1]$ 1-1 & onto

We get that $F(X_f)$ is uniform $[0,1]$

Thus: $u \lesssim f$ but we want $f \lesssim u$

Lets find F^{-1} ie

Solve for x in $y = F(x) = 1 - e^{-\beta x}$

iff $e^{-\beta x} = 1 - y$

iff $-\beta x = \ln(1-y)$

iff $x = -\frac{1}{\beta}\ln(1-y)$ but $1-y$ is uniform $[0,1]$

$$x = \frac{-\ln(X_u)}{\beta} = X_{Exp}$$

Thus $Exp \lesssim u$

Generating Normal Dist

The PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$

Setting $\sigma = 1$ we get Gauss's Unit Normal

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

What if we try computing CDF!

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

Thm $F(x)$ is not an elementary fcn!

note $f(x) = x e^{-x^2/2}$ is OK since

$$\frac{d}{dx}\left(-e^{-x^2/2}\right) = x e^{-x^2/2}$$

Trick: compute 2D normal.

Let $f(x,y) = \frac{1}{2\pi} e^{-x^2/2} \cdot e^{-y^2/2}$

$$= \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

$$= \frac{1}{2\pi} e^{-r^2/2} \text{ (in polar)}$$

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

Cumulative is prob in a disk of radius r.

$$D(R) = \int_0^R \frac{2\pi}{2\pi} r e^{-r^2/2} dr = \int_0^R r e^{-r^2/2} dr$$

$$-e^{-r^2/2}\Big]_0^R = 1 - e^{-R^2/2}$$

$Y = 1 - e^{-R^2/2}$	Alg u, v uniform $[0,1]$
$e^{-R^2/2} = 1 - Y$	in polar return (r, θ)
$-\frac{R^2}{2} = \ln(1-Y)$	where $r = \sqrt{-2\ln u}$
$R = \sqrt{-2\ln(1-Y)}$	$\theta = 2\pi v$

or return $x = r\cos\theta$

$y = r\sin\theta$

The Box-Muller Alg

Alg BM(u,v) where u & v are uniform [0,1].

1) Set $U = 2U-1$ & $V = 2V-1$ (uniform $[-1,1]$)

2) Set $W = U^2 + V^2$

   if $W > 1$ then restart.

   (U,V) is uniform over unit disk.

3) Set $A = \sqrt{\frac{-2\ln W}{W}}$ (scaling)

4) Return $T_1 = AU$ & $T_2 = (AV)$

Claim: BM generates 2D unit Gaussian.

Pf After step 2).

Write (U,V) as $V_1 = R\cos\theta$

$V_2 = R\sin\theta$

$S = R^2$

After step 4) we get (X_1, X_2)

$$X_1 = \sqrt{\frac{-2\ln S}{S}}\, V_1 = \sqrt{\frac{-2\ln S}{R^2}}\,(R\cos\theta) = \sqrt{-2\ln S}\,\cos\theta$$

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

In polar (R', θ') where $R' = \sqrt{-2\ln S}$, $\theta = [0, 2\pi]$

Lets compute CDF of R' ie

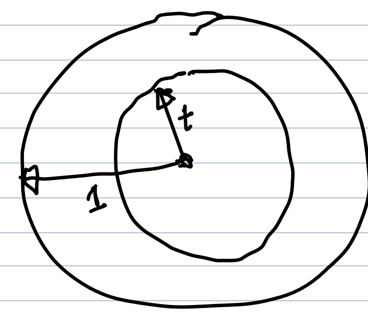
$$\begin{aligned} CDF(R') &= Pr[R' \le r] \\ &= Pr[\sqrt{-2\ln S} \le r] \\ &= Pr[-2\ln S \le r^2] \\ &= Pr[S \ge e^{-r^2/2}] \qquad (*) \end{aligned}$$

<u>Claim</u> $(*) = 1 - e^{-r^2/2}$

Suppose (u, v) is uniform over unit disk

i.e, PDF $= 1/4\pi$

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



Consider RV $S = \tilde{R}^2 = |u,v|^2 = u^2 + v^2$

$Prob[R^2 \le t] = Prob[R \le \sqrt{t}] = 1 - t$

this proves claim.