UC Berkeley
Department of Electrical Engineering and Computer Sciences

ELECTRICAL ENGINEERING 126: PROBABILITY AND RANDOM PROCESSES

Discussion 3
Fall 2017

---

1. **Order Statistics**

   For n a positive integer, let $X_1, \dots, X_n$ be i.i.d. random variables with common density function f and CDF F. For $i = 1, \dots, n$, let $X^{(i)}$ be the ith smallest of $X_1, \dots, X_n$, so we have $X^{(1)} \le \cdots \le X^{(n)}$. $X^{(i)}$ is known as the **ith order statistic**.

   (a) What is the CDF of $X^{(i)}$?

   (b) Differentiate the CDF to obtain the density of $X^{(i)}$.

   (c) Can you obtain the density of $X^{(i)}$ directly?

   Solution:

   (a) Observe that $\mathbb{P}(X^{(i)} \le x)$ is the probability that at least i of $X_1, \dots, X_n$ are $\le x$. We can split up the event $\{X^{(i)} \le x\}$ into the disjoint union of the events $C_k(x)$, $k = i, \dots, n$, where

   $$C_k(x) = \{\text{exactly } k \text{ of } X_1, \dots, X_n \text{ are} \le x\}.$$

   Thus,

   $$\mathbb{P}(X^{(i)} \le x) = \sum_{k=i}^{n} \mathbb{P}\big(C_k(x)\big) = \sum_{k=i}^{n} \binom{n}{k} F(x)^k \big(1 - F(x)\big)^{n-k}.$$

   (b) We differentiate the CDF to obtain

   $$\begin{aligned}
   f_{X^{(i)}}(x) &= \frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=i}^{n} \binom{n}{k} F(x)^k \big(1 - F(x)\big)^{n-k} \\
   &= \sum_{k=i}^{n} \binom{n}{k} f(x) F(x)^{k-1} \big(1 - F(x)\big)^{n-k-1} \\
   &\qquad \times \Big(k\big(1 - F(x)\big) - (n-k)F(x)\Big) \\
   &= \sum_{k=i}^{n} \binom{n}{k} k f(x) F(x)^{k-1} \big(1 - F(x)\big)^{n-k} \\
   &\qquad - \sum_{k=i}^{n-1} \binom{n}{k} (n-k) f(x) F(x)^k \big(1 - F(x)\big)^{n-k-1} \\
   &= \sum_{k=i}^{n} n \binom{n-1}{k-1} f(x) F(x)^{k-1} \big(1 - F(x)\big)^{n-k}
   \end{aligned}$$

$$
\begin{aligned}
&\quad - \sum_{k=i}^{n-1} n\binom{n-1}{k} f(x)F(x)^k\big(1-F(x)\big)^{n-k-1} \\
&= \sum_{k=i}^{n} n\binom{n-1}{k-1} f(x)F(x)^{k-1}\big(1-F(x)\big)^{n-k} \\
&\quad - \sum_{k=i+1}^{n} n\binom{n-1}{k-1} f(x)F(x)^{k-1}\big(1-F(x)\big)^{n-k} \\
&= n\binom{n-1}{i-1} f(x)F(x)^{i-1}\big(1-F(x)\big)^{n-i}.
\end{aligned}
$$

(c) The probability that $X^{(i)}$ lives in the small interval $[x, x+\varepsilon]$ is the product of:

- $nf(x)\varepsilon$, since there are n ways to choose which point is the ith largest and $f(x)\varepsilon$ is the probability that the random variable lies in the interval $[x, x+\varepsilon]$;
- $\binom{n-1}{i-1}F(x)^{i-1}$, because there are $\binom{n-1}{i-1}$ ways to choose which $i-1$ points will be smaller than $X^{(i)}$, and $F(x)^{i-1}$ is the probability that these points will be $\leq x$;
- $(1-F(x))^{n-i}$, as the other $n-i$ points must be found $\geq x$.

Hence, the desired probability is $n\binom{n-1}{i-1}f(x)F(x)^{i-1}(1-F(x))^{n-i}\varepsilon$, so the density is

$$
f_{X^{(i)}}(x) = n\binom{n-1}{i-1} f(x)F(x)^{i-1}\big(1-F(x)\big)^{n-i}.
$$

2. **Change of Variables**

(a) Suppose that X has the **standard normal distribution**, that is, X is a continuous random variable with density function

$$
f(x) = \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big).
$$

What is the density function of $\exp X$? (The answer is called the **log-normal distribution**.)

(b) Suppose that X is a continuous random variable with density f. What is the density of X^2?

(c) What is the answer to the previous question when X has the standard normal distribution? (This is known as the **chi-squared distribution**.)

Solution:

(a) We observe that for $x > 0$,

$$
\mathbb{P}(\exp X \leq x) = \mathbb{P}(X \leq \ln x) = F(\ln x),
$$

where F is the CDF of the standard normal distribution. So,

$$\begin{aligned} f_{\exp X}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}(\exp X \le x) = \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}(X \le \ln x) = F'(\ln x) \cdot \frac{1}{x} \\ &= \frac{f(\ln x)}{x} = \frac{1}{\sqrt{2\pi}x}\exp\Big(-\frac{(\ln x)^2}{2}\Big), \qquad x > 0. \end{aligned}$$

(b) We have $\mathbb{P}(X^2 \le x) = \mathbb{P}(-\sqrt{x} \le X \le \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f(x)\,\mathrm{d}x$. So, the density of X^2 is

$$f_{X^2}(x) = \frac{1}{2\sqrt{x}}\big(f(-\sqrt{x}) + f(\sqrt{x})\big).$$

(c) When X has the standard normal distribution,

$$f_{X^2}(x) = \frac{1}{2\sqrt{x}}\frac{1}{\sqrt{2\pi}}\Bigg(\exp\Big(-\frac{x}{2}\Big) + \exp\Big(-\frac{x}{2}\Big)\Bigg) = \frac{1}{\sqrt{2\pi x}}\exp\Big(-\frac{x}{2}\Big).$$

3. Expected Norm

Pick two points X and Y independently and uniformly in $[0,1]^2$. Calculate $\mathbb{E}[\|X - Y\|_2^2]$.

Solution:

If we let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$, then $X_1, X_2, Y_1, Y_2 \sim \text{Uniform}[0,1]$ and

$$\mathbb{E}[\|X - Y\|_2^2] = \mathbb{E}[(X_1 - Y_1)^2] + \mathbb{E}[(X_2 - Y_2)^2].$$

We can calculate $\mathbb{E}[(X_1 - Y_1)^2] = \mathbb{E}[X_1]^2 - 2\,\mathbb{E}[X_1]\,\mathbb{E}[Y_1] + \mathbb{E}[Y_1]^2 = 2/3 - 1/2 = 1/6$. So, $\mathbb{E}[\|X - Y\|_2^2] = 1/3$.