CS 536 : Homework Zero

16:198:536

1) Let $X_1, X_2, X_3, \ldots, X_n$ be i.i.d. random variables uniformly distributed on $[0, L]$. What is the density of

$$Y = \mathbf{max}\left[X_1, X_2, \ldots, X_n\right]? \tag{1}$$

To answer the question 'what is the distribution' of a random variable, frequently you want to look at what $\mathbb{P}(Y \leq x)$ is, as a function of x. This is the cumulative distribution function of Y, and we can recover the density by taking the derivative.

$$\mathbb{P}\left(Y \leq x\right) = \mathbb{P}\left(\max_i X_i \leq x\right) = \mathbb{P}\left(X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x\right). \tag{2}$$

By the assumption of independence, this factors:

$$\mathbb{P}\left(Y \leq x\right) = \mathbb{P}\left(X_1 \leq x\right)\mathbb{P}\left(X_2 \leq x\right)\ldots\mathbb{P}\left(X_n \leq x\right). \tag{3}$$

By the assumption that the X_i are identically distributed, this simplifies to

$$\mathbb{P}\left(Y \leq x\right) = \mathbb{P}\left(X \leq x\right)^n, \tag{4}$$

where X has the Unif$[0, L]$ distribution. The density of X is simply $1/L$ over the interval $[0, L]$, which means that the above probability is simply x/L. This simplifies to

$$\mathbb{P}\left(Y \leq x\right) = \left(\frac{x}{L}\right)^n. \tag{5}$$

Hence we have a c.d.f. of $F(x) = (x/L)^n$. Taking the derivative yields the density:

$$f(x) = n\frac{1}{L}\left(\frac{x}{L}\right)^{n-1}. \tag{6}$$

2) Let $X_1, X_2, X_3, \ldots, X_n$ be i.i.d. random variables with a normal distribution $N(\mu, \sigma^2)$.

- What is the joint density of these random variables?

The joint density of i.i.d. random variables is simply the product of the densities, i.e.,

$$f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}\left(x_i - \mu\right)^2\right). \tag{7}$$

Importantly, this simplifies to

$$\begin{aligned}
f(x_1, \ldots, x_n) &= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i - \mu\right)^2\right) \\
&= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i^2 - 2\mu x_i + \mu^2\right)\right) \\
&= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2\right]\right) \\
&= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{n}{2\sigma^2}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2\mu\bar{x}_n + \mu^2\right]\right) \\
&= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{n}{2\sigma^2}\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}_n^2\right) + \left(\bar{x}_n^2 - 2\mu\bar{x}_n + \mu^2\right)\right]\right) \\
&= \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}_n^2\right) - \frac{n}{2\sigma^2}\left(\bar{x}_n - \mu\right)^2\right)
\end{aligned} \tag{8}$$

With some additional algebra you can show that

$$\sum_{i=1}^{n} x_i^2 - n\bar{x}_n^2 = \sum_{i=1}^{n}\left(x_i - \bar{x}_n\right)^2, \tag{9}$$

hence

$$f(x_1, \ldots, x_n) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}\left(x_i - \bar{x}_n\right)^2\right) - \frac{n}{2\sigma^2}\left(\bar{x}_n - \mu\right)^2\right) \tag{10}$$

– Define the sample mean and sample variance as

$$\begin{aligned}
\bar{X}_n &= \frac{1}{n}\sum_{i=1}^{n} X_i \\
S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.
\end{aligned} \tag{11}$$

Show that $\bar{X}_n \sim N(\mu, \sigma^2/n)$, that $(n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1}$, and $\bar{X}_n$ and S_n^2 are independent.

The easiest of the three is showing that $\bar{X}_n \sim N(\mu, \sigma^2/n)$. $\bar{X}_n$ is a scaled sum of normal random variables, so is itself normal. The expected value and variance are given by

$$\begin{aligned} \mathbb{E}\left[\bar{X}_n\right] &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[X_i\right] = \frac{1}{n}n\mu = \mu \\ \mathrm{Var}(\bar{X}_n) &= \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{\sigma^2}{n}. \end{aligned} \tag{12}$$

That $(n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1}$ is a χ^2_{n-1}-random variable, and independent from $\bar{X}_n$, can be seen from the following equivalence:

$$-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 = -\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2\right) - \frac{n}{2\sigma^2}\left(\bar{X}_n-\mu\right)^2, \tag{13}$$

or, more expressively,

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}(n-1)S_n^2 + \left(\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}\right)^2. \tag{14}$$

Roughly, the LHS can be seen as the sum of n-many i.i.d. squared standard normals, or a χ^2_n-random variable. The RHS must have the same distribution, and is the sum of this Mystery Random Variable and the square of a standard normal. In order to balance the two sides, it must be that $(n-1)S_n^2/\sigma^2$ is distributed independently like the sum of $(n-1)$ i.i.d. squared standard normals, or as a χ^2_{n-1}-random variable, independent of $\bar{X}_n$. Note, a more precise proof relies on moment generating functions, but this gets the basic idea across.

3) Let (X_i, Y_i) for $i = 1, \ldots n$ be a set of points in $\mathbb{R}^2$. Find w^*, b^* to minimize

$$\sum_{i=1}^{n}\left(wX_i + b - Y_i\right)^2. \tag{15}$$

In the usual way, we want to solve for when the derivatives with respect to w, b are zero:

$$\begin{aligned}\sum_{i=1}^{n} 2X_i(wX_i + b - Y_i) &= 0 \\ \sum_{i=1}^{n} 2(wX_i + b - Y_i) &= 0.\end{aligned} \tag{16}$$

Expanding the summations and dividing through by n, this simplifies to

$$\begin{aligned}w\overline{X^2}_n + b\overline{X}_n - \overline{XY}_n &= 0 \\ w\overline{X}_n + b - \overline{Y}_n &= 0,\end{aligned} \tag{17}$$

which solves to

$$\begin{aligned}w^* &= \frac{\overline{XY}_n - \overline{X}_n\overline{Y}_n}{\overline{X^2}_n - \overline{X}_n^2} \\ b^* &= \frac{\overline{X^2}_n\overline{Y}_n - \overline{X}_n\overline{XY}_n}{\overline{X^2}_n - \overline{X}_n^2}.\end{aligned} \tag{18}$$

4) Continuing on the previous problem, suppose that for each (X_i, Y_i), we had that the X_i value is given, and $Y_i = wX_i + b + Z_i$ where $Z_i \sim N(0, \sigma^2)$.

   - What is the expected value of w^* in this case? What is the expected value of b^*?
   - What is the variance of w^*? What is the variance of b^*?

   Simplify as much as possible.

Extending off the previous problem, we now want to estimate the 'true' value of w and b based on the data. Using the formulae from the previous problem as the estimators:

$$\begin{aligned}\hat{w} &= \frac{\overline{XY}_n - \overline{X}_n\overline{Y}_n}{\overline{X^2}_n - \overline{X}_n^2} \\ &= \frac{\overline{wX^2 + bX + XZ}_n - \overline{X}_n\overline{wX + b + Z}_n}{\overline{X^2}_n - \overline{X}_n^2} \\ &= \frac{w\overline{X^2}_n + b\overline{X}_n + \overline{XZ}_n - \overline{X}_n\left(w\overline{X}_n + b + \overline{Z}_n\right)}{\overline{X^2}_n - \overline{X}_n^2} \\ &= w + \frac{\overline{XZ}_n - \overline{X}_n\overline{Z}_n}{\overline{X^2}_n - \overline{X}_n^2} \\ &= w + \frac{1}{\overline{X^2}_n - \overline{X}_n^2}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)Z_i\right)\end{aligned} \tag{19}$$

From the above form (a sum of scaled normals is normal), it can be shown that $\hat{w} \sim N(w, \sigma^2/n)$, and hence functions as a good estimator for w. The estimator for b can be analyzed similarly.

5) Let (A, B), (C, D) be two points in $\mathbb{R}^2$ chosen such that A, B, C, D are all i.i.d. standard normal random variables. Consider the line drawn from (A, B) to (C, D). What is the distribution of its length, and what is the expected value of the length?

The length of the line will be given by

$$\sqrt{(A-C)^2+(B-D)^2}. \tag{20}$$

Since A, C are i.i.d. standard normals, $A - C \sim N(0, 2)$, which we can express as $\sqrt{2}Z_1$ where Z_1 is a standard normal. Similarly, $B - D \sim \sqrt{2}Z_2$, where Z_1, Z_2 are independent. Substituting in:

$$\sqrt{(A-C)^2+(B-D)^2} = \sqrt{2Z_1^2+2Z_2^2} = \sqrt{2}\sqrt{Z_1^2+Z_2^2}. \tag{21}$$

But then $Z_1^2 + Z_2^2$ is a χ^2-random variable with two degrees of freedom. Hence:

$$\mathbb{E}\left[\sqrt{(A-C)^2+(B-D)^2}\right] = \int_0^\infty \sqrt{2}\sqrt{x}\frac{1}{2}e^{-x/2}dx = \sqrt{\pi}. \tag{22}$$