MTHE/STAT 353 – Solutions: Assignment 2

Winter, 2017

1. [Ghahramani, 9.1, #20]

Since \( \min(X_1, \ldots, X_n) \) is a nonnegative random variable, for any \( x > 0 \) we

\[
P(Y_n > x) = P(n \cdot \min(X_1, \ldots, X_n) > x) = P\left( \min(X_1, \ldots, X_n) > \frac{x}{n} \right) = \prod_{i=1}^{n} P\left( X_i > \frac{x}{n} \right) = \prod_{i=1}^{n} \left( 1 - F\left( \frac{x}{n} \right) \right)
\]

where \( F \) denotes the common cdf of the \( X_i \) and the second equality holds by independence. When \( n > x \), we have \( 0 < x/n < 1 \), and \( F(x/n) = x/n \), so

\[
\lim_{n \to \infty} P(Y_n > x) = \lim_{n \to \infty} \left( 1 - \frac{x}{n} \right)^n = e^{-x}.
\]

2. [Let \( X_1, X_3, X_2, X_4 \) be independent. . .]

According to the hint, the event that we cannot form a 4-sided quadrilateral is given by

\[
A^c = \{ X_1 > X_2 + X_3 + X_4 \} \cup \{ X_2 > X_1 + X_3 + X_4 \} \\
\cup \{ X_3 > X_1 + X_2 + X_4 \} \cup \{ X_4 > X_1 + X_2 + X_3 \}
\]

(i.e., it is the event that one of the rods is longer than the sum of the lengths of the other 3 rods). The 4 events above in the union comprising \( A^c \) are disjoint and, by symmetry, equally likely. Therefore

\[
P(A^c) = 4P(X_1 > X_2 + X_3 + X_4).
\]

Integrating the joint density of \( (X_1, X_2, X_3, X_4) \) over the region

\[
\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : 0 < x_i < 1 \text{ for } i = 1, 2, 3, 4 \text{ and } x_1 > x_2 + x_3 + x_4 \}
\]

we obtain

\[
P(A^c) = 4 \int_{0}^{1} \int_{0}^{1} \int_{0}^{x_1-x_2} \int_{0}^{x_1-x_2-x_3} dx_4 dx_3 dx_2 dx_1
\]
\[
\begin{align*}
&= 4 \int_0^1 \int_0^{x_1} \int_0^{x_1-x_2} (x_1-x_2-x_3) \, dx_3 \, dx_2 \, dx_1 \\
&= 4 \int_0^1 \int_0^{x_1} \left[ (x_1-x_2)x_3 - \frac{x_3^2}{2} \right] \, dx_2 \, dx_1 \\
&= 4 \int_0^1 \int_0^{x_1} \frac{(x_1-x_2)^2}{2} \, dx_2 \, dx_1 \\
&= 4 \int_0^1 \left[ -\frac{(x_1-x_2)^3}{6} \right]^{x_1} \, dx_1 \\
&= 4 \int_0^1 \frac{x_1^3}{6} \, dx_1 \\
&= 4 \frac{1}{24} = \frac{1}{6}
\end{align*}
\]

Therefore,
\[ P(A) = 1 - P(A^c) = 1 - \frac{1}{6} = \frac{5}{6}. \]

3. [Let \( X_1, \ldots, X_n \) be independent…]

(a) Define the event \( A = \{X_1 < X_2 < \cdots < X_{n-1} < X_n\} \). Then \( A = \{(X_1, \ldots, X_n) \in B\} \), where
\[ B = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 < x_2 < \cdots < x_{n-1} < x_n\}. \]

Letting \( \hat{f} \) denote the joint pdf of \( (X_1, \ldots, X_n) \), we have
\[ P(X_1 < X_2 < \cdots < X_{n-1} < X_n) = \int_B \cdots \int \hat{f}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n. \]

If \( f \) is continuous and positive on \( \mathbb{R} \), then \( F(x) = \int_{-\infty}^x f(t) \, dt \) is a continuously differentiable and strictly increasing (and thus invertible) function. Its inverse \( F^{-1} \) maps \((0,1)\) onto \( \mathbb{R} \) and is also continuously differentiable. We obtain that the mapping \( h : (x_1, \ldots, x_n) \mapsto (y_1, \ldots, y_n) \) given by
\[ y_i = h_i(x_i) = F(x_i), \quad i = 1, \ldots, n \]
is continuously differentiable having a continuously differentiable inverse \( g \) given by
\[ x_i = g_i(y_i) = F^{-1}(y_i), \quad i = 1, \ldots, n. \]

By multidimensional change of variables
\[ \int_B \cdots \int \hat{f}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n = \int_{h(B)} \cdots \int |J_g(y_1, \ldots, y_n)| \, dy_1 \cdots dy_n \]
where \( h(B) = \{ h(x_1, \ldots, x_n) : (x_1, \ldots, x_n) \in B \} \) is the image of \( B \) under the mapping \( h \) and \( J_g \) is the Jacobian of \( g \). But since \( u < v \) if and only if \( F(u) < F(v) \) and since \( F(\mathbb{R}) = (0,1) \), we have
\[
h(B) = \{ (y_1, \ldots, y_n) : 0 < y_1 < y_2 < \cdots < y_{n-1} < y_n < 1 \}.
\]

Also, since \( \frac{\partial}{\partial y_j} g_i = 0 \) if \( i \neq j \) and
\[
\frac{\partial}{\partial y_i} g_i = \frac{\partial}{\partial y_i} (F^{-1})(y_i) = \frac{1}{f(F^{-1}(y_i))}
\]
we have
\[
J_g(y_1, \ldots, y_n) = \prod_{i=1}^{n} \frac{1}{f(F^{-1}(y_i))}.
\]

By independence, we have \( \hat{f}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i) \), so we obtain
\[
P(X_1 < X_2 < \cdots < X_{n-1} < X_n) = \int_{B} \cdots \int_{B} \hat{f}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n
\]
\[
= \int_{h(B)} \cdots \int_{h(B)} \hat{f}(g_1(y_1), \ldots, g(x_n)) |J_g(y_1, \ldots, y_n)| \, dy_1 \cdots dy_n
\]
\[
= \int_{h(B)} \cdots \int_{h(B)} \prod_{i=1}^{n} f(F^{-1}(y_i)) \frac{1}{\prod_{i=1}^{n} f(F^{-1}(y_i))} \, dy_1 \cdots dy_n
\]
\[
= \int_{h(B)} \cdots \int_{h(B)} dy_1 \cdots dy_n
\]
\[
= \int_{0}^{1} \int_{0}^{y_n} \int_{0}^{y_{n-1}} \cdots \int_{0}^{y_2} dy_1 dy_2 \cdots dy_{n-2} dy_{n-1} dy_n
\]

which indeed does not depend on \( F \) (or \( f \)).

(b)
\[
p = \int_{0}^{1} \int_{0}^{y_n} \int_{0}^{y_{n-1}} \cdots \int_{0}^{y_2} dy_1 dy_2 \cdots dy_{n-2} dy_{n-1} dy_n
\]
\[
= \int_{0}^{1} \int_{0}^{y_n} \int_{0}^{y_{n-1}} \cdots \int_{0}^{y_2} dy_2 \cdots dy_{n-2} dy_{n-1} dy_n
\]
\[
= \int_{0}^{1} \int_{0}^{y_n} \int_{0}^{y_{n-1}} \cdots \int_{0}^{\frac{y_3}{2}} dy_3 \cdots dy_{n-2} dy_{n-1} dy_n
\]
\[
= \int_{0}^{1} \int_{0}^{y_n} \int_{0}^{y_{n-1}} \cdots \int_{0}^{\frac{y_4}{6}} dy_4 \cdots dy_{n-2} dy_{n-1} dy_n
\]
\[
= \cdots
\[
\int_{0}^{1} \int_{0}^{y_n \frac{y_n^{n-2}}{(n-2)!} \, dy_n \, dy_{n-1} = \int_{0}^{1} \frac{y_n^{n-1}}{(n-1)!} \, dy_n = \left[ \frac{y_n^n}{n!} \right]_0^1 = \frac{1}{n!}
\]

The intuition behind this result is the following: For any permutation \( \sigma \) of \( \{1, \ldots, n\} \), the event

\[A_\sigma = \{X_{\sigma(1)} < X_{\sigma(2)} < \cdots < X_{\sigma(n)}\}\]

has the same probability as the event \( \{X_1 < X_2 < \cdots < X_n\} \) since the independence and identical distribution of the \( X_i \) imply that \( (X_{\sigma(1)}, X_{\sigma(2)}, \ldots, X_{\sigma(n)}) \) has the same distribution as \( (X_1, X_2, \ldots, X_n) \). Also, since \( X_1, \ldots, X_n \) are jointly continuous, the union of the disjoint events \( A_\sigma \) has probability 1. Since there are \( n! \) such permutations, each of the events \( A_\sigma \) has probability \( 1/n! \), so \( P(X_1 < X_2 < \cdots < X_n) = 1/n! \).

4. [Ghahramani, 9.2, #3]

We have that \( X_1, \ldots, X_4 \) are independent and exponentially distributed with parameter \( \lambda \). By Theorem 9.5 (or class notes) the distribution function of \( X_{(4)} = \max(X_1, X_2, X_3, X_4) \) is

\[
F_4(x) = \binom{4}{4} F(x)^4 = \begin{cases}
(1 - e^{-\lambda x})^4 & \text{for } x > 0 \\
0 & \text{for } x \leq 0.
\end{cases}
\]

The desired probability is

\[
P(X_{(4)} \geq 3\lambda) = 1 - F_4(3\lambda) = 1 - (1 - e^{-3\lambda^2})^4.
\]

5. [Ghahramani, 9.2, #6]

From Theorem 9.6 (or class notes), the joint pdf of \( X_{(1)} \) and \( X_{(n)} \), the minimum and the maximum of the sample \( X_1, \ldots, X_n \), is given by

\[
f_{1n}(x, y) = \begin{cases}
n(n-1)f(x)f(y)(F(y) - F(x))^{n-2} & \text{for } x < y \\
0 & \text{otherwise},
\end{cases}
\]

where \( f(\cdot) \) and \( F(\cdot) \) are the pdf and cdf of \( X_i \). If we let \( M = (X_{(1)} + X_{(n)})/2 \) denote the midrange of the sample, then the distribution function of \( M \), denoted by \( G(t) \) is computed as

\[
G(t) = P(M \leq t) = P\left(\frac{X_{(1)} + X_{(n)}}{2} \leq t\right) = P(X_{(1)} + X_{(n)} \leq 2t).
\]
Therefore, we can compute $G(t)$ by integrating $f_{1n}(x, y)$ over the region (of the $xy$-plane) satisfying $x + y \leq 2t$ and $x \leq y$, as indicated in the hint. As $(x, y)$ varies over this region, $x$ varies between $-\infty$ and $t$ ($x$ cannot be larger than $t$ since $y \geq x$ would imply $x + y > 2t$ if $x$ were larger than $t$). For a given value of $x \in (-\infty, t]$, $y$ can range from $x$ to $2t - x$. Thus,

$$G(t) = \int_{-\infty}^{t} \int_{x}^{2t-x} f_{1n}(x, y) dy dx$$

$$= \int_{-\infty}^{t} \int_{x}^{2t-x} n(n - 1)f(x)f(y)(F(y) - F(x))^{n-2} dy dx$$

$$= \int_{-\infty}^{t} \left[ nf(x)(F(y) - F(x))^{n-1} \right]^{2t-x}_{x} dx$$

$$= \int_{-\infty}^{t} nf(x)(2t - x) - F(x) f(x) dx$$

as desired.

6. [Let $X_1, \ldots, X_n$ be independent…]

(a) By Theorem 9.6, the joint density of $X_{(1)}$ and $X_{(n)}$ is

$$f_{1n}(x, y) = n(n - 1)(F(y) - F(x))^{n-2} f(x)f(y),$$

for $x < y$ (and $f_{1n}(x, y) = 0$ for $x \geq y$). We will compute the distribution function of $R_n$ by computing $P(R_n \leq r)$ for any $r > 0$ and then differentiate this with respect to $r$ to get the probability density function of $R_n$. Let $r > 0$ be given. We compute

$$P(R_n \leq r) = P(X_{(n)} - X_{(1)} \leq r) = \int \int f_{1n}(x, y) dy dx$$

$$= \int_{-\infty}^{\infty} \int_{x}^{x+r} n(n - 1)(F(y) - F(x))^{n-2} f(x)f(y) dy dx$$

$$= \int_{-\infty}^{\infty} \left[ n(F(y) - F(x))^{n-1} f(x) \right]^{x+r}_{x} dx$$

$$= \int_{-\infty}^{\infty} n(F(x + r) - F(x))^{n-1} f(x) dx.$$

Differentiating this with respect to $r$ we obtain

$$f_{R_n}(r) = \frac{d}{dr} P(R_n \leq r) = \frac{d}{dr} \int_{-\infty}^{\infty} n(F(x + r) - F(x))^{n-1} f(x) dx$$

$$= \int_{-\infty}^{\infty} \frac{d}{dr} n(F(x + r) - F(x))^{n-1} f(x) dx$$

$$= \int_{-\infty}^{\infty} n(n - 1)(F(x + r) - F(x))^{n-2} f(x + r)f(x) dx$$
\begin{align*}
    &= n(n-1) \int_{-\infty}^{\infty} (F(x+r) - F(x))^{n-2} f(x+r)f(x)dx.
\end{align*}

Clearly, \( P(R_n \leq r) = 0 \) for \( r \leq 0 \), so \( f_{R_n}(r) = 0 \) for \( r \leq 0 \).

(b) When \( f \) and \( F \) are the Exponential(\( \lambda \)) probability density function and probability distribution function, respectively, we may compute \( f_{R_n} \) explicitly as

\begin{align*}
    f_{R_n}(r) &= n(n-1) \int_{0}^{\infty} (e^{-\lambda u} - e^{-\lambda(u+r)})^{n-2} \lambda e^{-\lambda(u+r)} \lambda e^{-\lambda u} du \\
    &= n(n-1) \int_{0}^{\infty} e^{-\lambda u(n-2)} (1 - e^{-\lambda r})^{n-2} \lambda^2 e^{-2\lambda u} e^{-\lambda r} du \\
    &= \lambda(n-1)(1 - e^{-\lambda r})^{n-2} \lambda e^{-\lambda r} \\
    &= \lambda(n-1)(1 - e^{-\lambda r})^{n-2} e^{-\lambda r},
\end{align*}

for \( r > 0 \), and \( f_{R_n}(r) = 0 \) for \( r \leq 0 \).