Tutorial 1

Definitions and Notation

- Let $X_1, \ldots, X_n$ be independent identically distributed. Such a collection is a *random sample*. The *order statistics* are defined as: $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$; $X_{k:n}$ is the kth lowest of the collection. Another common notation is the following: the order statistics of a sample size n are often written as $X_{(1)}, \ldots, X_{(n)}$
- The notation $X \sim Be(p)$ will be used to denote a Bernoulli trial with success probability p; $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$.
- The notation $Y \sim Bi(n, p)$ will be used to denote the Binomial distribution;

$$\mathbb{P}(Y = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & k = 0, 1, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

- Let X be a random variable. If Y is a continuous random variable with density f_Y, then

$$\mathbb{P}(X \in A) = \int \mathbb{P}(X \in A | Y = y) f_Y(y) dy.$$

Exercises

1. A generalisation of i.i.d. random variables is *exchangeable* random variables, an idea due to de Finetti (1972). The random variables $X_1, \ldots, X_n$ are *exchangeable* if any permutation of any subset of them of size $k \leq n$ has the same distribution. This exercise gives an example of random variables that are exchangeable but not i.i.d. Let $X_i|P$ be i.i.d. Be(P) variables, where $P \sim U(0, 1)$. That is, P is uniformly distributed on $[0, 1]$. Conditioned on $P = p$, the variables $X_1, \ldots, X_n$ are i.i.d. $Be(p)$ variables.

   (a) Show that the marginal distribution of any k of the Xs is the same as

   $$\mathbb{P}(X_1 = x_1, \ldots, X_k = x_k) = \int_0^1 p^t (1-p)^{k-t} dp = \frac{t!(k-t)!}{(k+1)!},$$

   where $t = \sum_{i=1}^k x_i$.

   (b) Show that, marginally,

   $$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) \neq \prod_{j=1}^n \mathbb{P}(X_j = x_j).$$

   The distribution is exchangeable, but not i.i.d.

2. Let $U_i : 1, 2, \ldots$ be i.i.d. $U(0, 1)$ variables. (uniformly distributed on the interval $[0, 1]$).

   (a) Consider a random sample of size n, $U_1, \ldots, U_n$. Compute $p_{U_{1:n}}$, the density for the first order statistic.

(b) Suppose that X has the distribution

$$\mathbb{P}(X = x) = \frac{c}{x!} \qquad x = 1, 2, 3, \ldots$$

where $c = 1/(e-1)$. Find the distribution of

$$Z = \min\{U_1, \ldots, U_X\}.$$

Hint Condition on X.

3. Let $X_1, \ldots, X_n$ be a random sample from a population with p.d.f. (probability density function)

$$p_X(x) = \frac{1}{\theta}\mathbf{1}_{[0,\theta]}(x)$$

Let $X_{1:n} < \ldots < X_{1:n}$ be the *order statistics*.

(a) Show that the joint density of $(X_{1:n}, \ldots, X_{n:n})$ is:

$$p_{X_{1:n},\ldots,X_{n:n}}(x_1, \ldots, x_n) = \frac{n!}{\theta^n}\mathbf{1}(0 \le x_1 \le \ldots \le x_n \le \theta).$$

(b) Show that $\frac{X_{1:n}}{X_{n:n}}$ and $X_{n:n}$ are independent random variables.

4. Let $X_1, \ldots, X_n$ be a random sample from a population with p.d.f.

$$p_X(x) = \frac{a}{\theta^a}x^{a-1}\mathbf{1}_{[0,\theta]}(x).$$

Let $X_{1:n} < \ldots < X_{n:n}$ be the order statistics. Show that $\frac{X_{1:n}}{X_{2:n}}, \frac{X_{2:n}}{X_{3:n}}, \ldots, \frac{X_{n-1:n}}{X_{n:n}}$ and $X_{n:n}$ are mutually independent random variables. Find the distribution of each of them.

5. Let X be a random variable with continuous distribution function F. Find the distribution function of $F(X)$.

6. Let $U \sim \text{Unif}(0,1)$ (that is, U has uniform distribution over the interval $[0,1]$) and let F be a continuous, monotonically increasing cumulative distribution function. Show that the c.d.f. of $Y := F^{-1}(U)$ is F. Is the result true if F is the c.d.f. of a discrete random variable, which takes values in the set $(x_j)_{j\ge1}$, $x_1 < x_2 < \ldots$?

7. Let $\{X_1, \ldots, X_n\}$ be a random sample from a population with continuous c.d.f. $F(x) = \mathbb{P}(X \le x)$ (that is $X_1, \ldots, X_n$ are independent identically distributed). Let

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{1}_{(-\infty,x]}(X_j).$$

$\widehat{F}_n$ is the *empirical distribution function.* Here $\mathbf{1}_A$ denotes the indicator function of a set A. That is,

$$\mathbf{1}_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases}$$

(a) Show that $n\widehat{F}_n(x) \sim Bi(n, F(x))$.

(b) Compute $\text{Cov}(\widehat{F}_n(x), \widehat{F}_n(y))$ (Cov denotes covariance).

(c) Let $D_n = \sup_{-\infty<x<+\infty} |\widehat{F}_n(x) - F(x)|$. Prove that the distribution of D_n is the same for all underlying continuous distribution functions F.

8. Let X and Y be independent random variables, with $\text{Gamma}(\alpha, \lambda)$ and $\text{Gamma}(\beta, \lambda)$ distributions respectively where $\alpha, \beta, \lambda > 0$. Show that $X + Y \sim \text{Gamma}(\alpha + \beta, \lambda)$. The $\text{Gamma}(\alpha, \lambda)$ distribution has density:

$$p(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbf{1}_{(0,+\infty)}(x).$$

9. Let $Z \sim N(0, 1)$ (standard normal). Find the distribution of Z^2. This is the $\chi^2(1)$ distribution.

10. Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$. Find the distribution of:

$$\frac{1}{\sigma^2} \sum_{k=1}^{n} (X_k - \mu)^2.$$

11. Let $W_1, \ldots, W_n$ be i.i.d. $\chi^2(1)$ random variables, then $W_1 + \ldots + W_n \sim \chi^2(n)$. Let Z and Y be independent random variables with standard normal and $\chi^2(n)$ distributions respectively. Find the density function of:

$$T = \frac{Z}{\sqrt{Y/n}}.$$

Short Answers

1. (a) Since they're conditionally independent given $P = p$, it follows that, for a sequence $(x_1, \ldots, x_k)$ with t success and $k-t$ failure,

$$\begin{aligned}\mathbb{P}(X_1 = x_1, \ldots, X_k = x_k) &= \int_0^1 \mathbb{P}(X_1 = x_1, \ldots, X_k = x_k | P = p) dp \\ &= \int_0^1 \prod_{j=1}^{k} \mathbb{P}(X_j = x_j | P = p) dp \\ &= \int_0^1 p^t (1-p)^{k-t} dp\end{aligned}$$

This is a standard integral, the beta integral. One way to compute it is by induction;

$$I(k,t) = \int_0^1 p^t (1-p)^{k-t} dp = \frac{k-t}{1+t} \int_0^1 p^{t+1} (1-p)^{k-t-1} dp = \frac{k-t}{1+t} I(k, t+1)$$

and

$$I(k,k) = \int_0^1 p^k dp = \frac{1}{k+1}$$

giving

$$I(k,t) = \frac{t!(k-t)!}{(k+1)!}.$$

(b)

$$\mathbb{P}(X_j = x_j) = \int_0^1 p^{x_j} (1-p)^{1-x_j} dp = \frac{1}{2} \qquad x_j = 0, 1$$

$$\prod_{j=1}^{n} \mathbb{P}(X_j = x_j) = \frac{1}{2^k} \qquad \{x_1, \ldots, x_k\} \in \{0,1\}^k$$

$$\mathbb{P}(X_1 = x_1, \ldots, X_k = x_k) = \frac{t!(k-t)!}{(k+1)!}$$

so they are not equal, hence the random variables are not i.i.d..

2. (a)

$$\mathbb{P}(U_{1:n} > x) = \mathbb{P}(U_1 > x, \ldots, U_n > x) = \mathbb{P}(U > x)^n = (1-x)^n \qquad 0 \le x \le 1$$

$$p_{U_{1:n}}(x) = n(1-x)^{n-1} \qquad 0 \le x \le 1$$

(b)

$$\mathbb{P}(Z > y) = \sum_{x=1}^{\infty} \mathbb{P}(U_{x:n} > y) \frac{c}{x!} = \sum_{x=1}^{\infty} \frac{c(1-y)^x}{x!} = c\left(e^{1-y} - 1\right) = \frac{e^{1-y} - 1}{e - 1} \qquad 0 \le y \le 1$$

$$F_Z(y) = 1 - \frac{e^{1-y} - 1}{e - 1} = \frac{e(1 - e^{-y})}{e - 1} \qquad 0 \le y \le 1.$$

3. (a) For any $A \subset \{0 < x_1 < \ldots < x_n \leq \theta\}$, let σ denote a permutation of $(1, \ldots, n)$ and let $A_\sigma = \{(x_1, \ldots, x_n) : (x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \in A\}$. By construction, the regions A_σ are disjoint.

$$\mathbb{P}((X_{1:n}, \ldots, X_{n:n}) \in A) = \sum_\sigma \mathbb{P}((X_1, \ldots, X_n) \in A_\sigma) = n!\mathbb{P}((X_1, \ldots, X_n) \in A)$$

so that, for all $A \subset \{0 < x_1 < \ldots < x_n \leq \theta\}$,

$$\mathbb{P}((X_{1:n}, \ldots, X_{n:n}) \in A) = \frac{n!}{\theta^n} \int \mathbf{1}_{\{0 \leq x_1 < \ldots < x_n \leq \theta\} \cap A} dx_1 \ldots dx_n$$

From this, it follows directly that the density is:

$$p_{X_{1:n}, \ldots, X_{1:n}}(x_1, \ldots, x_n) = \frac{n!}{\theta^n} \mathbf{1}(0 < x_1 < \ldots < x_n < \theta).$$

(b) Density of $(X_{1:n}, X_{n:n})$ computed as follows:

$$p_{X_{1:n}, X_{n:n}}(x_1, x_n) = \frac{n!}{\theta^n} \int_{x_1 < \ldots < x_n} \int dx_2 \ldots dx_{n-1} = \frac{n!}{\theta^n} \frac{1}{(n-2)!} (x_n - x_1)^{n-2}$$

(the integral is the area of the $n-2$ dimensional simplex) so

$$p_{X_{1:n}, X_{n:n}}(x_1, x_n) = \frac{n(n-1)}{\theta^n} (x_n - x_1)^{n-2} \qquad 0 < x_1 < x_n < \theta$$

Let $(Y, Z) = (\frac{X_{1:n}}{X_{n:n}}, X_{n:n})$ so that

$$(X_{1:n}, X_{n:n}) = (YZ, Z) \qquad Z = X_{n:n}.$$

Then

$$J_{(x_1, x_n) \to (y, z)} = \begin{pmatrix} z & 0 \\ y & 1 \end{pmatrix}$$

so the determinant is $|J_{(x_1, x_n) \to (y, z)}| = z$. It follows that

$$\begin{aligned} f_{(Y,Z)}(y, z) &= \frac{n(n-1)}{\theta^n} z (z - zy)^{n-2} \\ &= \frac{n(n-1)}{\theta^n} z^{n-1} (1 - y)^{n-2} \qquad 0 < yz < z < \theta \Rightarrow 0 < y < 1, 0 < z < \theta \end{aligned}$$

Density is in product form, therefore $Y \perp Z$.

$$p_Z(z) = \frac{n}{\theta^n} z^{n-1} \qquad 0 \leq z \leq \theta$$

$$p_Y(y) = (n-1)(1-y)^{n-2} \qquad 0 \leq y \leq 1$$

4.

$$p_{X_{1:n}, \ldots, X_{n:n}}(x_1, \ldots, x_n) = \frac{n! a^n}{\theta^{na}} (x_1 \ldots x_n)^{a-1} \mathbf{1}(0 \leq x_1 \leq \ldots \leq x_n \leq \theta).$$

$$(Y_1, \ldots Y_{n-1}, Y_n) = \left(\frac{X_{1:n}}{X_{2:n}}, \ldots, \frac{X_{n:n}}{X_{n-1:n}}, X_{n:n} \right)$$

$$(x_1,\ldots,x_n) = (\prod_{j=1}^n y_j, \prod_{j=2}^n y_j, \ldots, y_n)$$

(here $x_k = \prod_{j=k}^n y_j$).

$$J_{(x_1,\ldots,x_n)\to(y_1,\ldots,y_n)} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \ldots & \frac{\partial x_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial y_1} & \ldots & \frac{\partial x_n}{\partial y_n} \end{pmatrix} = \begin{pmatrix} \prod_{j=2}^n y_j & \prod_{j\neq 2} y_j & \ldots & \prod_{j=1}^{n-1} y_j \\ 0 & \prod_{j=3}^n y_j & \ldots & \prod_{j=2}^{n-1} y_j \\ \vdots & & \vdots & \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

This matrix is upper triangular, with determinant $\prod_{j=2}^n y_j^{j-1}$.

Density for $Y_1,\ldots,Y_{n-1},Y_n$ is therefore

$$p_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n) = \left(\prod_{j=2}^n y_j^{j-1}\right) \frac{n!a^n}{\theta^{na}} \left(\prod_{j=1}^n y_j^{j(a-1)}\right) \qquad 0 \le \prod_{j=1}^n y_j \le \prod_{j=2}^n y_j \le y_{n-1}y_n \le y_n \le \theta$$

which reduces to

$$p_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n) = \prod_{j=1}^n f_{Y_j}(y_j)$$

where

$$p_{Y_1}(y_1) = a y_1^{a-1} \qquad 0 \le y_1 \le 1$$
$$p_{Y_j}(y_j) = ja y_j^{ja-1} \qquad 0 \le y_j \le 1 \qquad 2 \le j \le n-1.$$
$$p_{Y_n}(y_n) = \frac{na}{\theta^{na}} y_n^{na-1} \qquad 0 \le y_n \le \theta.$$

5. The function F is continuous and non-decreasing. Define:

$$F^{-1}(u) = \sup\{x : F(x) < u\}.$$

Then $F(F^{-1}(u)) = u$, so that:

$$\mathbb{P}(F(X) \le x) = \mathbb{P}(X \le F^{-1}(x)) = F(F^{-1}(x)) = x \quad x \in [0,1]$$

hence $F(X) \sim \text{Unif}(0,1)$.

6.

$$\mathbb{P}(Y \le x) = \mathbb{P}(F^{-1}(U) \le x) = \mathbb{P}(U \le F(x)) = x.$$

Yes: define

$$F^{-1}(u) = \sup\{x : F(x) < u\}$$

if $(x_j)_{j\ge 1}$ denotes the points where the probability mass function has positive mass, where $\mathbb{P}(X = x_j) = p_j$ where $0 < x_1 < x_2 < \ldots$ and $\sum_j p_j = 1$, then

$$\begin{cases} \mathbb{P}(F^{-1}(U) \le x_1) = p_1 \\ \mathbb{P}(F^{-1}(U) \le x_{j+1}) - \mathbb{P}(F^{-1}(U) \le x_j) = p_{j+1} \qquad j \ge 1. \end{cases}$$

7. (a) $n\widehat{F}_n(x) = \sum_{j=1}^n \mathbf{1}_{\{X_j \le x\}}$. $n\widehat{F}_n(x)$ is therefore the number of 'success' in n Bernoulli trials each with success parameter $F(x)$, so $n\widehat{F}_n(x) \sim Bi(n, F(x))$. It follows that

$$\mathbb{E}\left[\widehat{F}_n(x)\right] = F(x) \qquad \text{Var}\left(\widehat{F}_n(x)\right) = \frac{F(x)(1-F(x))}{n}$$

(b)

$$\begin{aligned}
\text{Cov}(\widehat{F}_n(x), \widehat{F}_n(y)) &= \text{Cov}\left(\frac{1}{n}\sum_{j=1}^n \mathbf{1}(X_j \le x), \frac{1}{n}\sum_{k=1}^n \mathbf{1}(X_k \le y)\right) \\
&= \frac{1}{n^2}\sum_{j,k} \text{Cov}(\mathbf{1}(X_j \le x), \mathbf{1}(X_k \le y)) = \frac{1}{n^2}\sum_{j=1}^n \text{Cov}(\mathbf{1}(X_j \le x), \mathbf{1}(X_j \le y)) \\
&= \frac{1}{n}\left(\mathbb{E}\left[\mathbf{1}(X_j \le x)\mathbf{1}(X_j \le y)\right] - F(x)F(y)\right) \\
&= \frac{1}{n}F(x)(1-F(y)) \qquad x < y
\end{aligned}$$

(c) First consider F *strictly* increasing. Then

$$D_n = \sup_{-\infty<y<+\infty} |\widehat{F}_n(y) - F(y)| = \sup_{0<x<1} |\widehat{F}_n(F^{-1}(x)) - F(F^{-1}(x))| = \sup_{0<x<1} |\widehat{F}_n(F^{-1}(x)) - x|.$$

$$\widehat{F}_n(F^{-1}(x)) = \frac{1}{n}\sum_{j=1}^n \mathbf{1}_{\{X_j \le F^{-1}(x)\}} = \frac{1}{n}\sum_{j=1}^n \mathbf{1}_{\{F(X_j) \le x\}}$$

and the result follows since $F(X_1), \dots, F(X_n)$ are i.i.d. $U(0,1)$ variables.

General case: use $F^{-1*}(x) = \sup\{y | F(x) < y\}$. Then $F^{-1*}(x) \le z \Leftrightarrow x \le F(z)$ and same proof follows.

Therefore, for any continous F, the distribution of D_n is the same as that of

$$\sup_{0 \le y \le 1} |\widehat{F}_n(y) - y|$$

where

$$\widehat{F}_n(y) = \frac{1}{n}\sum_{j=1}^n \mathbf{1}_{\{U_j \le y\}}$$

for a random sample $U_1, \dots, U_n$ of $U(0,1)$ variables.

8. (8,9 and 10 form a sequence of questions, which follow on from each other). Let $W = X + Y$ then

$$F_W(w) = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^w x^{\alpha-1} e^{-\lambda x} \left(\int_0^{w-x} y^{\beta-1} e^{-\lambda y} dy\right) dx$$

so that

$$\begin{aligned}
p_W(w) &= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda w} \int_0^w x^{\alpha-1}(w-x)^{\beta-1} dx \\
&= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda w} w^{\alpha+\beta-1} \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx \propto w^{\alpha+\beta-1} e^{-\lambda w}.
\end{aligned}$$

Since this is a density function, it follows (using the formula for a Gamma density function given in the question) that

$$p_W(w) = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha+\beta)} w^{\alpha+\beta-1} e^{-\lambda w} \mathbf{1}_{(0,+\infty)}(w).$$

9. For $x > 0$,

$$\mathbb{P}(Z^2 \leq x) = \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

so

$$p_{Z^2}(x) = \frac{1}{\sqrt{2\pi x}} e^{-x/2} \mathbf{1}_{(0,+\infty)}(x) \propto x^{-1/2} e^{-x/2} \mathbf{1}_{(0,+\infty)}(x); \qquad \text{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right).$$

Note: the Gamma$(\frac{1}{2}, \frac{1}{2})$ is also known as a $\chi^2(1)$ distribution.

10. Let $Z_k = \frac{X_k - \mu}{\sigma}$ then $Z_1, \ldots, Z_n$ are i.i.d. $N(0,1)$ and $\frac{1}{\sigma^2} \sum_{k=1}^n (X_k - \mu)^2 = \sum_{j=1}^n Z_j^2$. Using the results of 8 and 9, therefore, this has distribution Gamma $\left(\frac{n}{2}, \frac{1}{2}\right)$.

This is also known as a $\chi^2(n)$ distribution.

11.

$$\mathbb{P}(T \leq x) = \mathbb{P}\left(Z \leq x\sqrt{\frac{Y}{n}}\right) = \frac{1}{\sqrt{2\pi}} \frac{(1/2)^{n/2}}{\Gamma(n/2)} \int_0^\infty \int_{-\infty}^{x\sqrt{\frac{y}{n}}} e^{-z^2/2} y^{(n/2)-1} e^{-y/2} dz dy.$$

The density is:

$$\begin{aligned}
p_T(x) &= \frac{1}{2^{(n+1)/2}\pi^{1/2}n^{1/2}\Gamma(n/2)} \int_0^\infty y^{(n-1)/2} \exp\left\{-\frac{y}{2}\left(1+\frac{x^2}{n}\right)\right\} dy \\
&= \frac{(1+\frac{x^2}{n})^{-(n+1)/2}}{2^{(n+1)/2}\pi^{1/2}n^{1/2}\Gamma(n/2)} \int_0^\infty v^{(n+1)/2-1} e^{-v/2} dv \\
&= \frac{\Gamma(\frac{n+1}{2})}{n^{1/2}\pi^{1/2}\Gamma(\frac{n}{2})} \frac{1}{(1+\frac{x^2}{n})^{(n+1)/2}}
\end{aligned}$$