1. This is to review second year probability material. We will use indicators and the
definition of independence from Karr’s book here (cf. slides 50, 54).

Look: independence of $A_1, \ldots, A_n$ is equivalent to that of $\mathbf{1}_{A_1}, \ldots, \mathbf{1}_{A_n}$. Likewise, independence of $A_1^c, \ldots, A_n^c$ is equivalent to that of $\mathbf{1}_{A_1^c}, \ldots, \mathbf{1}_{A_n^c}$.

Next, since

$$\{\mathbf{1}_{A_j^c} \in B_j\} = \{1 - \mathbf{1}_{A_j} \in B_j\} = \{\mathbf{1}_{A_j} \in B_j^c\}$$

with $B_j^c = 1 - B_j := \{1 - x : x \in B_j\}$, it now follows from the definition of independence of RVs that $\mathbf{1}_{A_1^c}, \ldots, \mathbf{1}_{A_n^c}$ are independent iff $\mathbf{1}_{A_1}, \ldots, \mathbf{1}_{A_n}$ are.

That’s it, right? Right.

2. The density of $X$ is $f(t) = F'(t) = ae^{-at}$, $t > 0$.

(a) Here $g(x) = \sqrt{x}$, $x \geq 0$, which has the inverse $h(y) = y^2$, $y \geq 0$. As $h'(y) = 2y$, we get

$$f_{Y_1}(y) = f(h(y))|h'(y)| = 2aye^{-ay^2}, \quad y > 0.$$ 

(b) Here $g(x) = x^2$, $x \geq 0$, which has the inverse $h(y) = \sqrt{y}$, $y \geq 0$. As $h'(y) = \frac{1}{2\sqrt{y}}$, we get

$$f_{Y_2}(y) = f(h(y))|h'(y)| = \frac{a}{2\sqrt{y}} e^{-a\sqrt{y}}, \quad y > 0.$$ 

(c) This is a discrete RV, its distribution is given by the values

$$P([X] = k) = P(X \in [k, k+1)) = F(k+1) - F(k) = (1 - e^{-a(k+1)}) - (1 - e^{-ak}) = (1 - e^{-a})e^{-ak}, \quad k = 0, 1, \ldots$$

(which is a geometric distribution, right?).

(d) First compute the DF: for $t \in [0, 1)$,

$$F_{Y_4}(t) = P(X - [X] \leq t) = \sum_{k=0}^{\infty} P(X \in [k, k+t]) = \sum_{k=0}^{\infty} (F(k+t) - F(k))$$

$$= \sum_{k=0}^{\infty} ((1 - e^{-a(k+t)}) - (1 - e^{-ak})) = \sum_{k=0}^{\infty} (1 - e^{-at})e^{-ak} = \frac{1 - e^{-at}}{1 - e^{-a}}$$

(this is the sum of a geometric series!). Hence the density is

$$f_{Y_4}(y) = F_{Y_4}'(y) = \frac{ae^{-ay}}{1 - e^{-a}}, \quad y \in (0, 1).$$

(e) Here $Y_5 = F(X)$, so must have $Y_5 \sim U(0, 1)$.

Computing the distribution of $Y_5$ directly: here $g(x) = 1 - e^{-ax}$, $x \geq 0$, which has the inverse $h(y) = -\frac{1}{a} \ln(1 - y)$, $y \in (0, 1)$. As $h'(y) = \frac{1}{a(1-y)}$, we get

$$f_{Y_5}(y) = f(h(y))|h'(y)| = \frac{a}{a(1-y)} \exp\{-a(-\frac{1}{a} \ln(1 - y))\} = 1, \quad y \in (0, 1).$$
3. We are good in computing the probabilities of intersections of independent events, right? So, in all the questions below, we’ll be trying to reduce the resp. problems to calculating such probabilities (e.g., by switching to the complements of unions etc.).

(a) As the desired event is a union of independent events, we have:

\[
F_{X(1)}(t) = P(X(1) \leq t) = P\left(\bigcup_{j \leq n} \{X_j \leq t\}\right) = 1 - \prod_{j \leq n} P(X_j > t) = 1 - P(X_1 > t)^n = 1 - (1 - F(t))^n.
\]

(b) Straightforward: \(F_{X(1)}(t) = P\left(\bigcap_{j \leq n} \{X_j \leq t\}\right) = \prod_{j \leq n} P(X_j \leq t) = F(t)^n.\)

(c) Here need to think a bit first:

\[
F_{X(1), X(n)}(t_1, t_2) = P(X(1) \leq t_1, X(n) \leq t_2) = P(X(n) \leq t_2) - P(X(1) > t_1, X(n) \leq t_2)
= F(t_2)^n - P\left(\bigcap_{j \leq n} \{t_1 < X_j \leq t_2\}\right) = F(t_2)^n - P(t_1 < X_1 \leq t_2)^n
= F(t_2)^n - (F(t_2) - F(t_1))^n.
\]

(d) One can argue as follows\(^1\). Fix an \(x \in \mathbb{R}\). For a small \(\Delta\), introduce events \(A := \{X_k \in [x, x + \Delta]\}\) (aiming to compute \(f_{X(k)}(x)\) as \(\lim_{\Delta \downarrow 0} P(A)/\Delta\)), \(B_1 := \\{\text{exactly one of the } X_j's \text{ is in } [x, x + \Delta]\}\), and \(B_2 := \{\text{at least two of the } X_j's \text{ are in } [x, x + \Delta]\}\). Then, as \(A \subset B_1 \cup B_2\) and \(B_1 B_2 = \emptyset\),

\[
P(A) = P(AB_1) + P(AB_2).
\]

Since \(B_2 \subset \bigcup_{i<j} C_i C_j\), where \(C_i := \{X_i \in (x, x + \Delta]\}\) has probability \(P(C_i) = F(x + \Delta) - F(x)\), we see that

\[
P(AB_2) \leq P(B_2) \leq \sum_{i<j} \frac{P(C_i C_j)}{P(C_i)P(C_j)} = \frac{n(n-1)}{2} (F(x + \Delta) - F(x))^2.
\]

As \(n\) is a fixed number, when dividing the last expression by \(\Delta\) and letting \(\Delta \downarrow 0\), we will get zero in the limit, right? (Note that \((F(x + \Delta) - F(x))/\Delta \to f(x)\) then.) So we only need to consider

\[
P(AB_1) = P\left(\begin{array}{c}k - 1 \text{ of the } X_j's \text{ are in } (-\infty, x]\; \text{ and } \; \text{ one of the } X_j's \text{ is in } (x, x + \Delta]\; \text{ and } \; n - k \text{ of the } X_j's \text{ are in } (x + \Delta, \infty]\end{array}\right)
= \frac{n!}{(k-1)!1!(n-k)!} F^{k-1}(x)(F(x + \Delta) - F(x))(1 - F(x + \Delta))^n-k,
\]

as this is the multinomial probability (NB: here \(n\) points are thrown at random, independently of each other, into three bins with “success” probabilities [of hitting the bins] \(F(x), F(x + \Delta) - F(x)\) and \(1 - F(x + \Delta)\), resp.) of the outcome \((k-1, 1, n-k)\) (cf. \url{http://en.wikipedia.org/wiki/Multinomial_distribution}).

---

\(^1\)The argument is slightly sloppy. To make it better, one should only consider those points \(x\) at which \(F'(x) = f(x) < \infty\). As that relation holds “almost everywhere” (i.e., on a set whose complement has zero length), we’ll be alright at the end of the day.
Therefore we obtain that, as $\Delta \to 0$,
\[
\frac{P(A)}{\Delta} = \frac{P(AB_1)}{\Delta} + \frac{P(AB_2)}{\Delta} \xrightarrow{\Delta \to 0} \frac{n!}{(k-1)!1!(n-k)!} F^{k-1}(x)f(x)(1 - F(x))^{n-k},
\]
the last expression giving the desired density of $X_{(k)}$.

(e) Here we take $x < y$ and small $\Delta > 0$ (s.t. $x + \Delta < y$) and, similarly to (d), use the relation
\[
P\{X(k) \in (x, x+\Delta], \, X(m) \in (y, y+\Delta]\} = f_{X(k),X(m)}(x,y)\Delta^2(1 + o(1)), \quad \Delta \to 0
\]
(which holds when $f_{X(k),X(m)}$ is continuous at $(x,y)$) to find the density $f_{X(k),X(m)}$.

First (again, similarly to (d)) one shows that the probability of more than two points hitting $(x, x+\Delta] \cup (y, y+\Delta]$ is $O(\Delta^3)$, and so the contribution of that event to the probability on the LHS on the above relation will be negligibly small when we’ll let $\Delta \downarrow 0$.

Hence should just compute the probability of the following intersection of events:

$k - 1$ of the $X_j$’s are in $(-\infty, x]$; one of the $X_j$’s is in $(x, x+\Delta]$; $m - k - 1$ of the $X_j$’s are in $(x + \Delta, y]$; one of the $X_j$’s is in $(y, y+\Delta]$; $n - m$ of the $X_j$’s are in $(x + \Delta, \infty]$

But this is again a multinomial probability! It’s equal to
\[
\frac{n!}{(k-1)!(m-k-1)!(n-m)!} F^{k-1}(x)(F(x+\Delta) - F(x)) \times (F(y) - F(x+\Delta))^{m-k-1}(F(y+\Delta) - F(y))(1 - F(y+\Delta))^{n-m}.
\]
Dividing by $\Delta^2$ and passing to the limit yields this expression for $f_{X(k),X(m)}$:
\[
\frac{n!}{(k-1)!(m-k-1)!(n-m)!} F^{k-1}(x)f(x)(F(y)-F(x))^{m-k-1}f(y)(1-F(y))^{n-m}.
\]