Back to Dirichlet Distribution

Properties of Dirichlet Distribution: Suppose $k \geq 2$ and

$$(Y_1, Y_2, \ldots, Y_{k-1}) \sim \text{Dir}(\alpha_1, \alpha_2, \ldots, \alpha_{k-1}; \alpha_k).$$

Then the following properties hold.

1. $(Y_1, Y_2, \ldots, Y_{k-2}) \sim \text{Dir}(\alpha_1, \alpha_2, \ldots, \alpha_{k-2}; \alpha_k + \alpha_{k-1})$ (provided $k \geq 3$).

   In fact, for all $l \in \{1, 2, \ldots, k-1\}$,

$$(Y_1, Y_2, \ldots, Y_l) \sim \text{Dir}\left(\alpha_1, \alpha_2, \ldots, \alpha_l; \sum_{l+1}^{k} \alpha_i\right).$$

Proof: Take $X_1, X_2, \ldots, X_k$ ind r.v.s such that each $X_i \sim \text{Gamma}(\alpha_i, 1)$. Then we know

$$\left(\frac{X_1}{S}, \frac{X_2}{S}, \ldots, \frac{X_{k-1}}{S}\right) \stackrel{d}{=} (Y_1, Y_2, \ldots, Y_{k-1}),$$

where $S = X_1 + X_2 + \cdots + X_k$. Fix $l \in \{1, 2, \ldots, k-1\}$.

Applying Thm 3 of Pg 244 with the projection map $T: \mathbb{R}^{k-1} \to \mathbb{R}^l$ defined by

$$T: (u_1, u_2, \ldots, u_{k-1}) \mapsto (u_1, u_2, \ldots, u_l)$$

we get $(Y_1, Y_2, \ldots, Y_l) \stackrel{d}{=} \left(\frac{X_1}{S}, \frac{X_2}{S}, \ldots, \frac{X_l}{S}\right)$.

Note that $X_1, X_2, \dots, X_l, \sum_{l+1}^{k} X_i$ are ind r.v.s with $X_1 \sim \text{Gamma}(\alpha_1, 1)$, $X_2 \sim \text{Gamma}(\alpha_2, 1), \dots,$ $X_l \sim \text{Gamma}(\alpha_l, 1)$ and $\sum_{l+1}^{k} X_i \sim \text{Gamma}\left(\sum_{l+1}^{k} \alpha_i, 1\right)$.

Of course, $S = X_1 + X_2 + \dots + X_l + \sum_{l+1}^{k} X_i$.

Therefore using the example in Pg 233 (more specifically, Remark 1 of Pg 241), we get

$$\left(\frac{X_1}{S}, \frac{X_2}{S}, \dots, \frac{X_l}{S}\right) \sim \text{Dir}\left(\alpha_1, \alpha_2, \dots, \alpha_l ; \sum_{l+1}^{k} \alpha_i\right),$$

from which it follows using Thm 2 of Pg 243 that $(Y_1, Y_2, \dots, Y_l) \sim \text{Dir}\left(\alpha_1, \alpha_2, \dots, \alpha_l ; \sum_{l+1}^{k} \alpha_i\right)$.

In order to use Thm 2, we apply the last observation of Pg 245, namely,

$$(Y_1, Y_2, \dots, Y_l) \stackrel{d}{=} \left(\frac{X_1}{S}, \frac{X_2}{S}, \dots, \frac{X_l}{S}\right).$$

This completes the proof of Property 1.

<u>Remark</u>: The above proof uses both Thm 2 and Thm 3 (see Pg 243 – 244) crucially.

2. For all permutation π of $\{1, 2, \dots, k-1\}$,

$$(Y_{\pi(1)}, \dots, Y_{\pi(k-1)}) \sim \text{Dir}(\alpha_{\pi(1)}, \dots, \alpha_{\pi(k-1)}; \alpha_k).$$

3. Any marginal is Dirichlet with appropriate parameters: $\forall$ distinct $i_1, i_2, \dots, i_l \in \{1, 2, \dots, k-1\}$ ($\Rightarrow l \leq k-1$), we have

$$(Y_{i_1}, Y_{i_2}, \dots, Y_{i_l}) \sim \text{Dir}\left(\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_l}; \sum_{j=1}^{k} \alpha_j - \sum_{p=1}^{l} \alpha_{i_p}\right).$$

4. $(Y_1 + Y_2, Y_3, Y_4, \dots, Y_{k-1}) \sim \text{Dir}(\alpha_1 + \alpha_2, \alpha_3, \alpha_4, \dots, \alpha_{k-1}; \alpha_k)$.

5. $Y_1 + Y_2 + \dots + Y_{k-1} \sim \text{Dir}\left(\sum_{i=1}^{k-1} \alpha_i ; \alpha_k\right) \equiv \text{Beta}\left(\sum_{1}^{k-1} \alpha_i, \alpha_k\right)$.

6. $\left(Y_1 + Y_2,\ Y_3 + Y_4 + Y_5 + Y_6,\ Y_7,\ \sum_{8}^{k-1} Y_i\right)$

$\sim \text{Dir}\left(\alpha_1 + \alpha_2,\ \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6,\ \alpha_7,\ \sum_{7}^{k-1} \alpha_i ;\ \alpha_k\right)$

(whenever $k \geq 9$).

7. Each $Y_i \sim \text{Dir}\left(\alpha_i ; \sum_{j=1}^{k} \alpha_j - \alpha_i\right) \equiv \text{Beta}\left(\alpha_i, \sum_{\substack{j=1 \\ j \neq i}}^{k} \alpha_j\right)$.

Exc: Prove Properties 2–7 above.

Exc: Write down "analogous properties" of multinomial dist^n and prove them.

Remark: Dirichlet distribution is useful as "prior distributions" for the multinomial parameters in Bayesian statistics. Properties (see, e.g., 1-7 in Pg 245 - 247) of Dirichlet distribution make it amenable to Bayesian statistics. ■

Slight Digression: Linear Algebra of nnd and pd matrices with real entries

Notations: (1) For $m, n \in \mathbb{N}$,

$$\mathbb{R}^{m\times n} := \{A : A \text{ is an } m\times n \text{ real matrix}\}$$

denotes the set of all $m\times n$ matrices with real entries.

(2) We shall identify, for each $m \in \mathbb{N}$, the set $\mathbb{R}^{m\times 1}$ of all column vectors of dimension m with $\mathbb{R}^m$.

Defn: A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is called nonnegative definite (nnd) if $\forall\ \underset{\sim}{x} \in \mathbb{R}^{n\times 1}$,

$$\underbrace{\underset{1\times n}{\underset{\sim}{x}^T}\ \underset{n\times n}{A}\ \underset{n\times 1}{\underset{\sim}{x}}}_{1\times 1} \geq 0.$$

e.g., ① I_n, λI_n with $\lambda \geq 0$, $\text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$ with $\alpha_1 \geq 0, \alpha_2 \geq 0, \dots, \alpha_n \geq 0$, etc.

② Take any $U \in \mathbb{R}^{n \times p}$ (for any $p \in \mathbb{N}$) and define $\underset{n\times n}{A} = \underset{n\times p}{U}\,\underset{p\times n}{U^T}$, then A is nnd.

Exc: Show that ② holds.

Remark: It can be shown that any symmetric nnd matrix arises in the fashion mentioned in ② above.

Defn: A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called positive definite (pd) if $\forall \underset{\sim}{x} \in \mathbb{R}^{n \times 1} \setminus \{\underset{\sim}{0}\}$,

$$\underset{\sim}{x}^T A \underset{\sim}{x} > 0.$$

Remark: Note that A is pd iff A is nnd and $\underset{\sim}{x}^T A \underset{\sim}{x} = 0 \Rightarrow \underset{\sim}{x} = \underset{\sim}{0}$. In particular, pd $\overset{\Rightarrow}{\not\Leftarrow}$ nnd.

e.g., ① I_n, λI_n with $\lambda > 0$, $\text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$ with $\alpha_1 > 0, \alpha_2 > 0, \dots, \alpha_n > 0$, etc.

② Take any nonsingular matrix $U \in \mathbb{R}^{n \times n}$ and define $A = UU^T$. Then A is a pd matrix.

Proof of ②: Clearly $A^T = (UU^T)^T = (U^T)^T U^T = UU^T = A$ yielding A is symmetric.

Take any $\underset{\sim}{x} \in \mathbb{R}^{n \times 1} \setminus \{\underset{\sim}{0}\}$. Then

$$\underset{\sim}{x}^T A \underset{\sim}{x} = \underset{\sim}{x}^T U U^T \underset{\sim}{x} \qquad [\because A = UU^T]$$

$$= (U^T \underset{\sim}{x})^T U^T \underset{\sim}{x}$$

$$= \underset{\sim}{y}^T \underset{\sim}{y}. \qquad \left[\text{Putting } \cancel{\underset{\sim}{y} = A\underset{\sim}{x}} \quad \underset{n \times n}{\underset{\sim}{y}} = \underset{n \times n}{U^T} \underset{n \times 1}{\underset{\sim}{x}}\right]$$

$\underset{\sim}{x} \neq \underset{\sim}{0}$ and U^T is nonsingular $(\because U$ is nonsing$)$

$$\Rightarrow \underset{(y_1, y_2, \dots, y_n)}{\underset{\parallel}{\underset{\sim}{y}}} = U^T \underset{\sim}{x} \neq \underset{\sim}{0}$$

$$\Rightarrow \underset{\sim}{x}^T A \underset{\sim}{x} = \underset{\sim}{y}^T \underset{\sim}{y} = \sum_{i=1}^{n} y_i^2 > 0$$

$\Rightarrow A$ is pd.

Thm: Any symmetric pd matrix $A \in \mathbb{R}^{n\times n}$ is of the form

$$\text{(PD)} \ldots \quad A = UU^T$$

for some nonsingular matrix $U \in \mathbb{R}^{n\times n}$.

Proof: See a book on Linear Algebra.

Remarks: ① (PD) can be used as the definition of a symmetric pd matrix with real entries.

② Note that if (PD) holds, then

$$\begin{aligned} \det(A) &= \det(UU^T) \\ &= \det(U)\det(U^T) \\ &= (\det(U))^2 \end{aligned}$$

$$\Rightarrow \det(U) = +\sqrt{\det(A)} \quad \text{or} \quad -\sqrt{\det(A)}$$

$$\Rightarrow |\det(U)| = +\sqrt{\det(A)}$$

This observation will be used later.