In particular, for all $(y_1, y_2, \ldots, y_{k-1}, s) \in J$,

$$h_1(y_1, y_2, \ldots, y_{k-1}, s) = y_1 s,$$

$$h_2(y_1, y_2, \ldots, y_{k-1}, s) = y_2 s,$$

$$\vdots$$

$$h_{k-1}(y_1, y_2, \ldots, y_{k-1}, s) = y_{k-1} s,$$

$$h_k(y_1, y_2, \ldots, y_{k-1}, s) = \left(1 - \sum_{i=1}^{k-1} y_i\right) s$$

so that $\quad g^{-1} = (h_1, h_2, \ldots, h_{k-1}, h_k)$ on J.

Exc: Check also that g is a bijection.

Therefore, we can compute the partial derivatives (on J): [all of them exist and are cont]

$$\frac{\partial h_1}{\partial y_1} = s, \quad \frac{\partial h_1}{\partial y_2} = 0, \quad \ldots, \quad \frac{\partial h_1}{\partial y_{k-1}} = 0, \quad \frac{\partial h_1}{\partial s} = y_1;$$

$$\frac{\partial h_2}{\partial y_1} = 0, \quad \frac{\partial h_2}{\partial y_2} = s, \quad \ldots, \quad \frac{\partial h_2}{\partial y_{k-1}} = 0, \quad \frac{\partial h_2}{\partial s} = y_2;$$

$$\vdots$$

$$\frac{\partial h_{k-1}}{\partial y_1} = 0, \quad \frac{\partial h_{k-1}}{\partial y_2} = 0, \quad \ldots, \quad \frac{\partial h_{k-1}}{\partial y_{k-1}} = s, \quad \frac{\partial h_{k-1}}{\partial s} = y_{k-1};$$

$$\frac{\partial h_k}{\partial y_1} = -s, \quad \frac{\partial h_k}{\partial y_2} = -s, \quad \ldots, \quad \frac{\partial h_k}{\partial y_{k-1}} = -s, \quad \frac{\partial h_k}{\partial s} = 1 - \sum_{i=1}^{k-1} y_i.$$

This leads to the following Jacobian matrix of the map g^{-1}:

- $J_{g^{-1}}(y_1, y_2, \ldots, y_{k-1}, s)$

$$= \begin{pmatrix} s & 0 & \cdots & 0 & y_1 \\ 0 & s & \cdots & 0 & y_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & s & y_{k-1} \\ -s & -s & \cdots & -s & 1-\sum_{i=1}^{k-1} y_i \end{pmatrix},$$

$(y_1, y_2, \ldots, y_{k-1}, s) \in J.$

Hence, on J its determinant is $\dfrac{d g^{-1}(y_1, \ldots, y_{k-1}, s)}{d(y_1, \ldots, y_{k-1}, s)}$

$$= \det\left(J_{g^{-1}}(y_1, y_2, \ldots, y_{k-1}, s)\right)$$

$$= \det \begin{pmatrix} s & 0 & \cdots & 0 & y_1 \\ 0 & s & \cdots & 0 & y_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & s & y_{k-1} \\ -s & -s & \cdots & -s & 1-\sum_{i=1}^{k-1} y_i \end{pmatrix}$$

$$= \det \begin{pmatrix} s & 0 & \cdots & 0 & y_1 \\ 0 & s & \cdots & 0 & y_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & s & y_{k-1} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} = s^{k-1} > 0$$

on J.

In particular, we can apply the change of multivariate joint density formula and obtain that $g(X_1, X_2, \ldots, X_k) = \left(\frac{X_1}{S}, \frac{X_2}{S}, \ldots, \frac{X_{k-1}}{S}, S\right)$ $= (Y_1, Y_2, \ldots, Y_{k-1}, S)$ is also a cont random vector with joint range J. Also for any $(y_1, y_2, \ldots, y_{k-1}, s) \in J$, a joint pdf of $(Y_1, Y_2, \ldots, Y_{k-1}, S)$ is given by

$$f_{Y_1, Y_2, \ldots, Y_{k-1}, S}(y_1, y_2, \ldots, y_{k-1}, s)$$

$$= f_{X_1, X_2, \ldots, X_k}\left(g^{-1}(y_1, y_2, \ldots, y_{k-1}, s)\right) \left| \frac{dg^{-1}(y_1, y_2, \ldots, y_{k-1}, s)}{d(y_1, y_2, \ldots, y_{k-1}, s)} \right|$$

$$= f_{X_1, X_2, \ldots, X_k}\left(y_1 s, y_2 s, \ldots, y_{k-1} s, \left(1 - \sum_{i=1}^{k-1} y_i\right) s\right) s^{k-1}.$$

By hypothesis, $X_1, X_2, \ldots, X_k$ are ind with each $X_i \sim \text{Gamma}(\alpha_i, \lambda)$. Thus a joint pdf of $(X_1, X_2, \ldots, X_k)$ is

$$f_{X_1, X_2, \ldots, X_k}(x_1, x_2, \ldots, x_k)$$

$$= \frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)} e^{-\lambda x_1} x_1^{\alpha_1 - 1} \cdot \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} e^{-\lambda x_2} x_2^{\alpha_2 - 1} \cdot \ldots \cdot \frac{\lambda^{\alpha_k}}{\Gamma(\alpha_k)} e^{-\lambda x_k} x_k^{\alpha_k - 1},$$

$$(x_1, x_2, \ldots, x_k) \in I = (0, \infty)^k.$$

$$= \frac{\lambda^{\alpha_1 + \alpha_2 + \cdots + \alpha_k}}{\Gamma(\alpha_1)\Gamma(\alpha_2)\ldots\Gamma(\alpha_k)} e^{-\lambda(x_1 + x_2 + \cdots + x_k)} x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} \ldots x_k^{\alpha_k - 1},$$

$$(x_1, x_2, \ldots, x_k) \in I = (0, \infty)^k.$$

Therefore, for any $(y_1, y_2, \ldots, y_{k-1}, s) \in J$,

$$f_{Y_1, Y_2, \ldots, Y_{k-1}, S}(y_1, y_2, \ldots, y_{k-1}, s)$$

$$= f_{X_1, X_2, \ldots, X_k}\left(y_1 s, y_2 s, \ldots, y_{k-1} s, \left(1 - \sum_{i=1}^{k-1} y_i\right) s\right) s^{k-1}$$

$$= \frac{\lambda^{\sum_1^k \alpha_i} s^{k-1}}{\prod_1^k \Gamma(\alpha_i)} e^{-\lambda s} (y_1 s)^{\alpha_1 - 1} \ldots (y_{k-1} s)^{\alpha_{k-1}} \left(\left(1 - \sum_1^{k-1} y_i\right) s\right)^{\alpha_k - 1}$$

$$= \frac{\lambda^{\sum_1^k \alpha_i}}{\prod_{i=1}^k \Gamma(\alpha_i)} e^{-\lambda s} s^{\sum_1^k \alpha_i - 1} y_1^{\alpha_1 - 1} y_2^{\alpha_2 - 1} \dots y_{k-1}^{\alpha_{k-1}-1} \left(1 - \sum_1^{k-1} y_i\right)^{\alpha_k - 1}.$$

Summarizing, we get that $(Y_1, Y_2, \dots, Y_{k-1}, S)$ is a cont random vector with a joint pdf

$$f_{Y_1, Y_2, \dots, Y_{k-1}, S}(y_1, y_2, \dots, y_{k-1}, s)$$

$$= \frac{\Gamma(\sum_1^k \alpha_i)}{\prod_1^k \Gamma(\alpha_i)} y_1^{\alpha_1 - 1} \dots y_{k-1}^{\alpha_{k-1}-1} \left(1 - \sum_1^{k-1} y_i\right)^{\alpha_k - 1} \cdot \frac{\lambda^{\sum_1^k \alpha_i}}{\Gamma(\sum_1^k \alpha_i)} e^{-\lambda s} s^{\sum_1^k \alpha_i - 1},$$

$$(y_1, y_2, \dots, y_{k-1}) \in (0,1)^{k-1}, \quad \sum_{i=1}^{k-1} y_i < 1, \quad s > 0$$

Clearly, $S \sim \text{Gamma}\left(\sum_1^k \alpha_i, \lambda\right)$,

$S \perp\!\!\!\perp (Y_1, Y_2, \dots, Y_{k-1})$, and

$(Y_1, Y_2, \dots, Y_{k-1})$ is a $(k-1)$-dimensional (recall that $k \geq 2 \Rightarrow k-1 \in \mathbb{N}$) cont random vector with a joint pdf given by

$$\cancel{f_{Y_1, Y_2, \dots, Y_{k-1}}(y_1, y_2, \dots, y_{k-1}) = \frac{\Gamma(\sum_1^k \alpha_i)}{\prod_1^k \Gamma(\alpha_i)} y_1^{\alpha_1 - 1}}$$

• $f_{Y_1, Y_2, \ldots, Y_{k-1}}(y_1, y_2, \ldots, y_{k-1})$

$$= \begin{cases} \dfrac{\Gamma(\sum_1^k \alpha_i)}{\prod_1^k \Gamma(\alpha_i)} y_1^{\alpha_1 - 1} y_2^{\alpha_2 - 1} \cdots y_{k-1}^{\alpha_{k-1}-1} \left(1 - \sum_1^{k-1} y_i\right)^{\alpha_k - 1} & \text{if } (y_1, y_2, \ldots, y_{k-1}) \in (0,1)^{k-1}, \ \sum_1^{k-1} y_i < 1, \\ 0 & \text{otherwise.} \end{cases} \quad \cdots (D)$$

__Def^n:__ Fix $k \geq 2$ and $\alpha_1, \alpha_2, \ldots, \alpha_k \in (0, \infty)$. A cont random vector $(Y_1, Y_2, \ldots, Y_{k-1})$ is said to follow a $(k-1)$-dimensional Dirichlet distribution with parameters $\alpha_1, \alpha_2, \ldots, \alpha_{k-1}$ and α_k if it has a joint pdf (D) as above.

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

__Remarks:__ ① We have just shown the following: if $k \geq 2$, $X_1, X_2, \ldots, X_k$ are ind r.v.s with each $X_i \sim \text{Gamma}(\alpha_i, \lambda)$, then

$$S := X_1 + X_2 + \cdots + X_k \sim \text{Gamma}\left(\sum_1^k \alpha_i, \lambda\right),$$

$$(Y_1, Y_2, \ldots, Y_{k-1}) := \left(\frac{X_1}{S}, \frac{X_2}{S}, \ldots, \frac{X_{k-1}}{S}\right) \sim \text{Dir}(\alpha_1, \ldots, \alpha_{k-1}; \alpha_k),$$

and

$$S \perp\!\!\!\perp (Y_1, Y_2, \ldots, Y_{k-1}) \mathrel{=:} \underset{\sim}{Y}.$$

In particular, we have shown that (D) is indeed a valid $(k-1)$-dimensional pdf.

② The independence of S and $\underset{\sim}{Y}$ ~~can~~ also follows from Basu's theorem. (see Pg 233)

3. For $k=2$, we already know (from the Remark on Pg 233)

that $Y_1 := \frac{X_1}{S} = \frac{X_1}{X_1+X_2} \sim \text{Beta}(\alpha_1, \alpha_2)$.

This means as distributions,

$$\text{Dir}(\alpha_1 ; \alpha_2) \equiv \text{Beta}(\alpha_1, \alpha_2),$$

i.e., $Y \sim \text{Dir}(\alpha_1 ; \alpha_2)$ iff $Y \sim \text{Beta}(\alpha_1, \alpha_2)$.

In particular, this means that the Dirichlet distn is a higher dimensional generalization of beta distribution. In other words, beta distn is the univariate Dirichlet distn.

4. Dirichlet distn is important in Bayesian statistics.

A Small Digression: Equality in Distribution

Defn: Two random vectors $(Y_1, Y_2, \ldots, Y_n) \doteqdot \underset{\sim}{Y}$ and $(Z_1, Z_2, \ldots, Z_n) \doteqdot \underset{\sim}{Z}$ are called equal in distribution if they have the same joint cdf, i.e., for all $(u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$,

$$P[Y_1 \leq u_1, Y_2 \leq u_2, \ldots, Y_n \leq u_n] = P[Z_1 \leq u_1, Z_2 \leq u_2, \ldots, Z_n \leq u_n].$$

Notation: $\underset{\sim}{Y} \overset{d}{=} \underset{\sim}{Z}$ or $(Y_1, Y_2, \ldots, Y_n) \overset{d}{=} (Z_1, Z_2, \ldots, Z_n)$.

Thm1: If $\underset{\sim}{Y} \overset{d}{=} \underset{\sim}{Z}$, then for any "nice" $A \subseteq \mathbb{R}^n$,

$$P[\underset{\sim}{Y} \in A] = P[(Y_1, \ldots, Y_n) \in A] = P[(Z_1, \ldots, Z_n) \in A] \cancel{= P[\underset{\sim}{Z} \in A]}$$
$$= P[\underset{\sim}{Z} \in A].$$

Proof: Beyond our scope.

Exc: If $U \sim \text{Unif}(0,1)$, then show that $(U, 1-U) \overset{d}{=} (1-U, U)$.

Thm 2: If $\underset{\sim}{Y} \overset{d}{=} \underset{\sim}{Z}$ and $\underset{\sim}{Y}$ is a cont random vector with a joint pdf φ, then $\underset{\sim}{Z}$ is also a cont random vector with a joint pdf φ.

Proof: For all $(u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$, we have

$$P(Z_1 \leq u_1, Z_2 \leq u_2, \ldots, Z_n \leq u_n)$$

$$= P[Y_1 \leq u_1, Y_2 \leq u_2, \ldots, Y_n \leq u_n] \quad [\text{~~Since~~} \because \underset{\sim}{Y} \overset{d}{=} \underset{\sim}{Z}]$$

$$= \int_{-\infty}^{u_n} \int_{-\infty}^{u_{n-1}} \cdots \int_{-\infty}^{u_1} \varphi(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n,$$

$$[\because \underset{\sim}{Y} \text{ has a joint pdf } \varphi]$$

which shows that $\underset{\sim}{Z}$ is also a cont r.v. with a joint pdf φ.

Remark: Thm 2 also holds if we replace "cont" by "discrete", and "pdf" by "pmf" everywhere.

Thm 3: If $\underset{\sim}{Y} \overset{d}{=} \underset{\sim}{Z}$ and $T: \mathbb{R}^n \to \mathbb{R}^m$ ($m \in \mathbb{N}$) is a "nice" map, then $T(\underset{\sim}{Y}) \overset{d}{=} T(\underset{\sim}{Z})$.

Proof: Take $(v_1, v_2, \dots, v_m) \in \mathbb{R}^m$. Then

$$P\Big[T(\underset{\sim}{Y}) \in \underbrace{(-\infty, v_1] \times (-\infty, v_2] \times \cdots \times (-\infty, v_m]}_{\overset{..}{=} B}\Big]$$

$$= P\Big[\underset{\sim}{Y} \in T^{-1}(B)\Big]$$

$$= P\Big[\underset{\sim}{Z} \in T^{-1}(B)\Big] \qquad [\text{By Thm } 1]$$

$$= P\Big[T(\underset{\sim}{Z}) \in B\Big]$$

$$= P\Big[T(\underset{\sim}{Z}) \in (-\infty, v_1] \times (-\infty, v_2] \times \cdots \times (-\infty, v_m]\Big].$$

This shows that

$$T(\underset{\sim}{Y}) \overset{d}{=} T(\underset{\sim}{Z}).$$