Dirichlet distribution

Dirichlet distribution

theta and alpha are k dimensional vectors.

Dirichlet distributions are all for simplexes.

$$\mathrm{Dir}(\theta|\alpha) = \frac{1}{B(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$

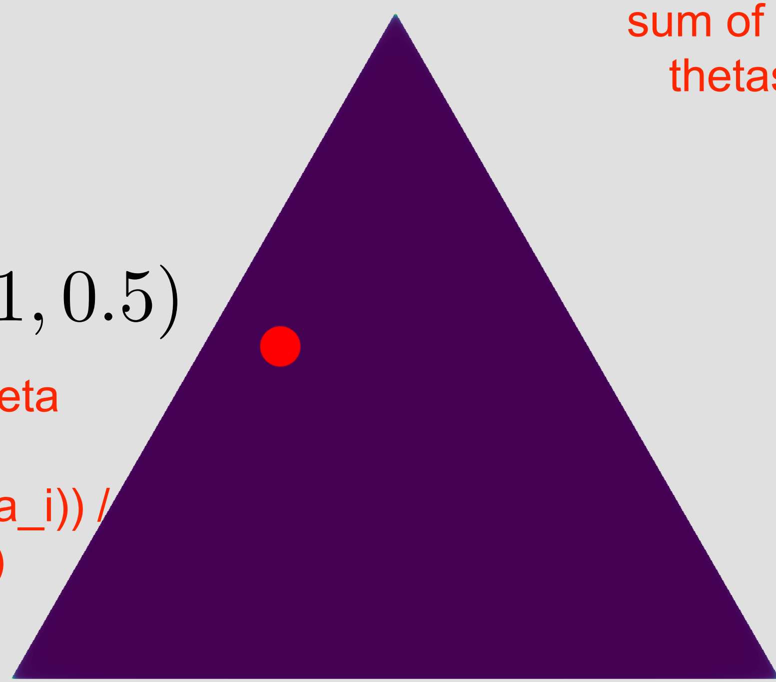
$$\sum_k \theta_k = 1$$

$$\theta_k \geq 0$$

^ This is a simplex.

a 2-simplex, is the set of points:
C = {theta0 mu0 + theta1mu1 + theta2mu2} such that the sum of all thetas is 1, and all thetas are non-negative.

$(0, 0, 1)$



$($ 0.4 $, 0.1, 0.5)$

$(1, 0, 0)$ $(0, 1, 0)$

NOTE: B(\alpha) is the beta function. Which is
B(a) = prod i to k (Gamma(a_i)) / gamma(sum i to k a_i)

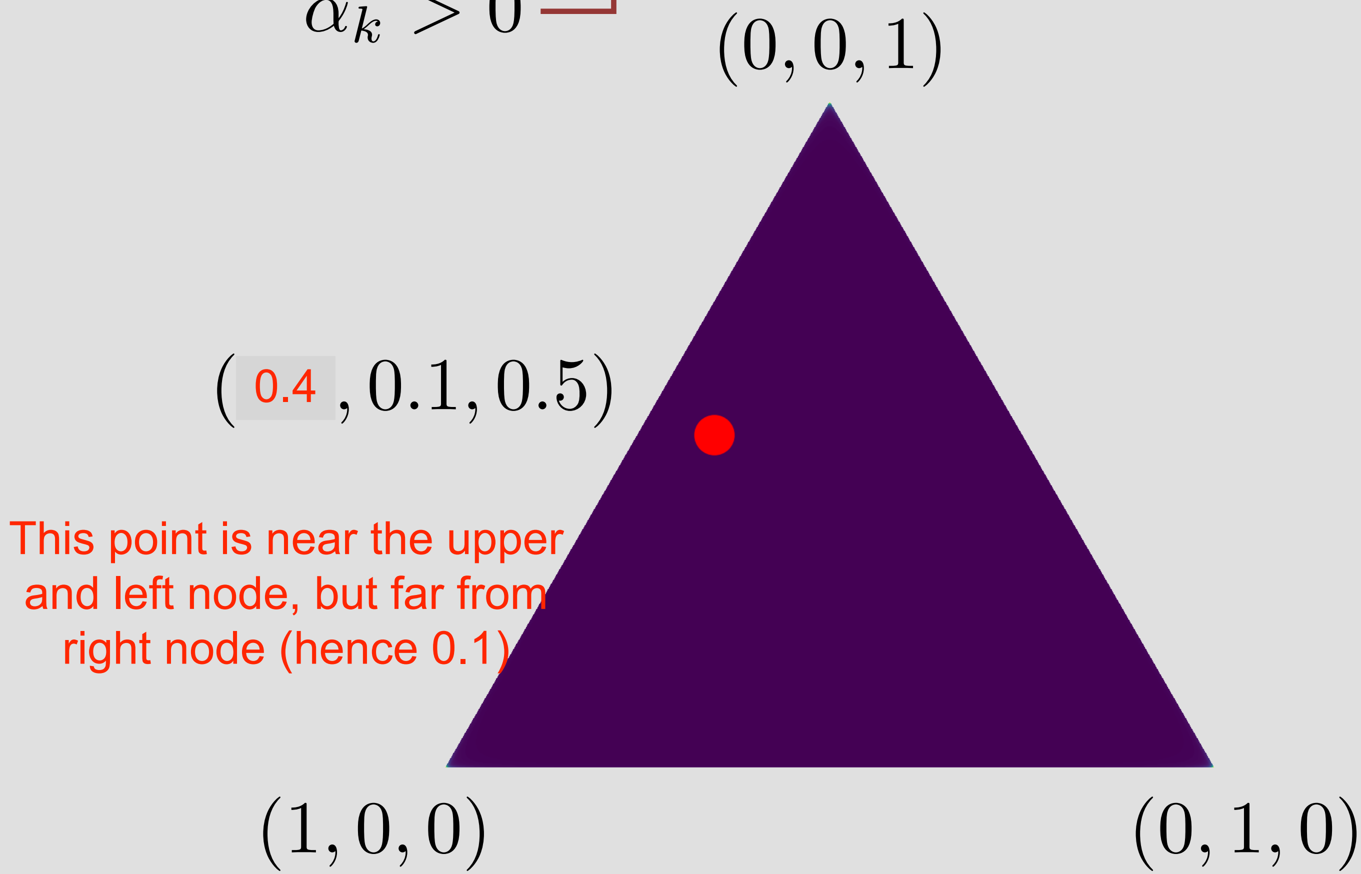
in general, an n-simplex the set of points C = {t0m0 + t1m1+ ... + tn mn}, where all t sum to 1, and are non-negative.

^NOTE: each mu here is a point in r^k.

Dirichlet distribution

$$\mathrm{Dir}(\theta|\alpha) = \frac{1}{B(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$

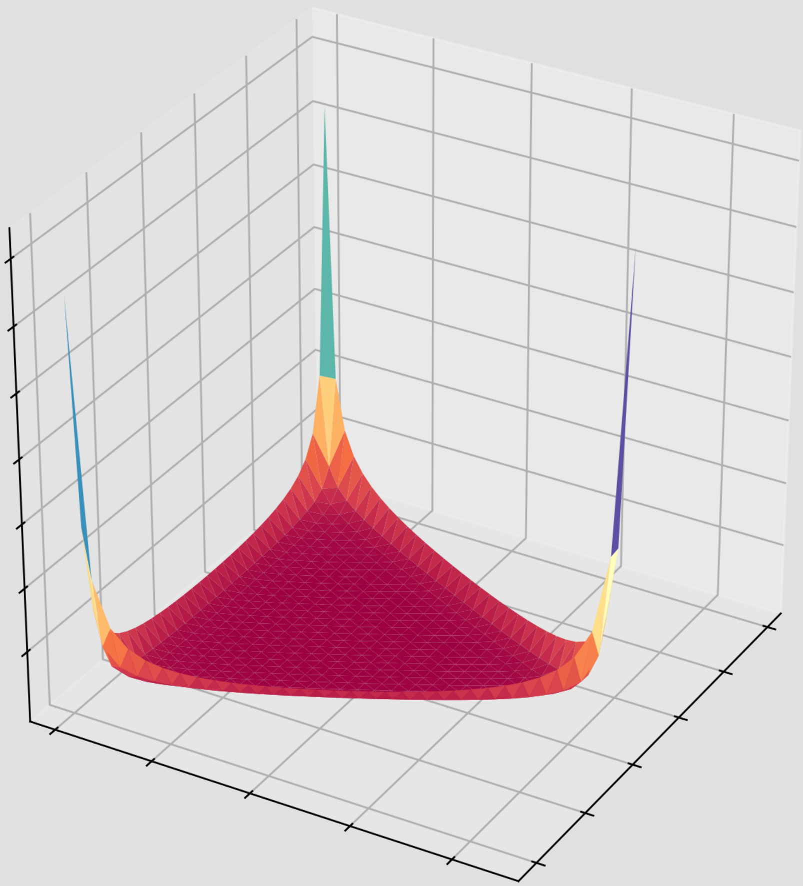
$\alpha_k > 0$



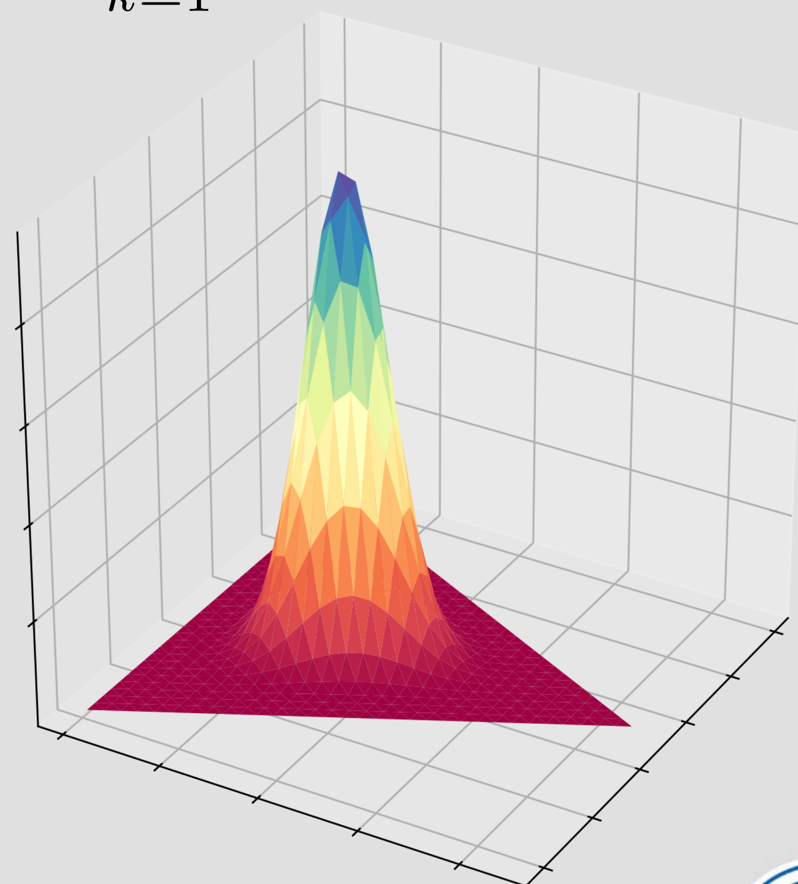
Dirichlet distribution

This is for a 2-simplex. Each point in the 2 simplex has a density as below.

$$\mathrm{Dir}(\theta|\alpha) = \frac{1}{B(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$



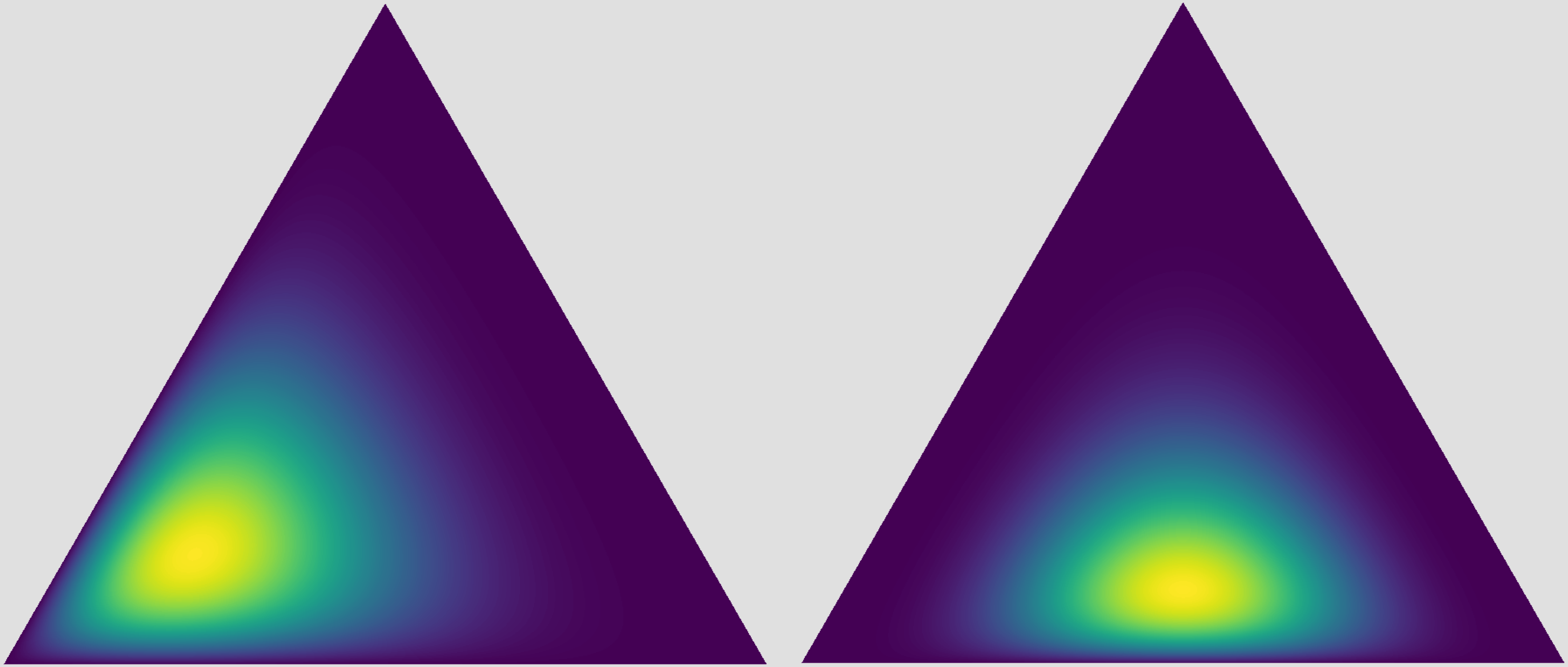
$$\alpha = (0.1, 0.1, 0.1)$$



$$\alpha = (10, 10, 10)$$

Dirichlet distribution

$$\mathrm{Dir}(\theta|\alpha) = \frac{1}{B(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$



$\alpha = (5, 2, 2)$ $\alpha = (5, 5, 2)$

alpha determines the concentration.

Statistics

$$\mathrm{Dir}(\theta|\alpha) = \frac{1}{B(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$

$$\mathbb{E}\theta_i = \frac{\alpha_i}{\alpha_0}$$

This is for a SPECIFIC theta component.

$$\mathrm{Cov}(\theta_i, \theta_j) = \frac{\alpha_i \alpha_0 [i=j] - \alpha_i \alpha_j}{\alpha_0^2(\alpha_0 + 1)}$$

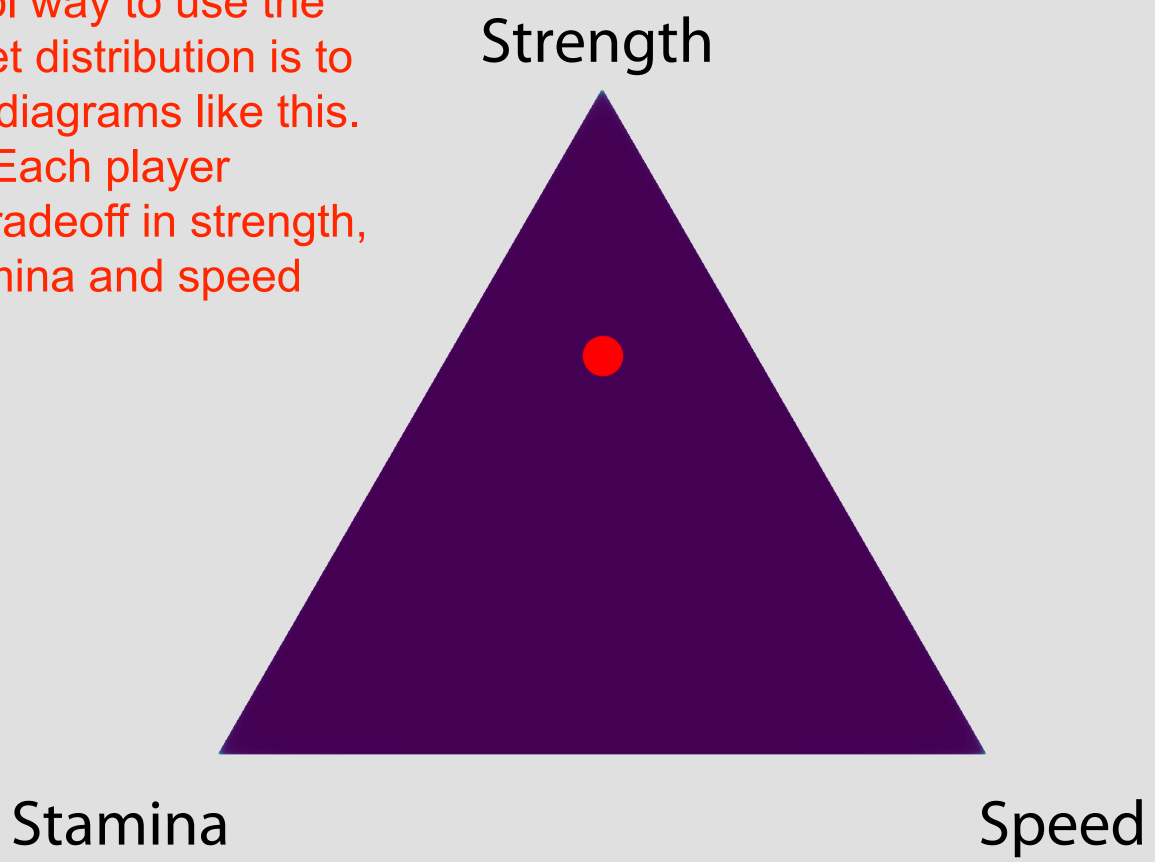
$$\alpha_0 = \sum_{k=1}^{K} \alpha_k$$

Example

Massively multiplayer online role-playing game (**MMORPG**)

Player 1:

A cool way to use the dirichlet distribution is to model diagrams like this. Each player has a tradeoff in strength, stamina and speed

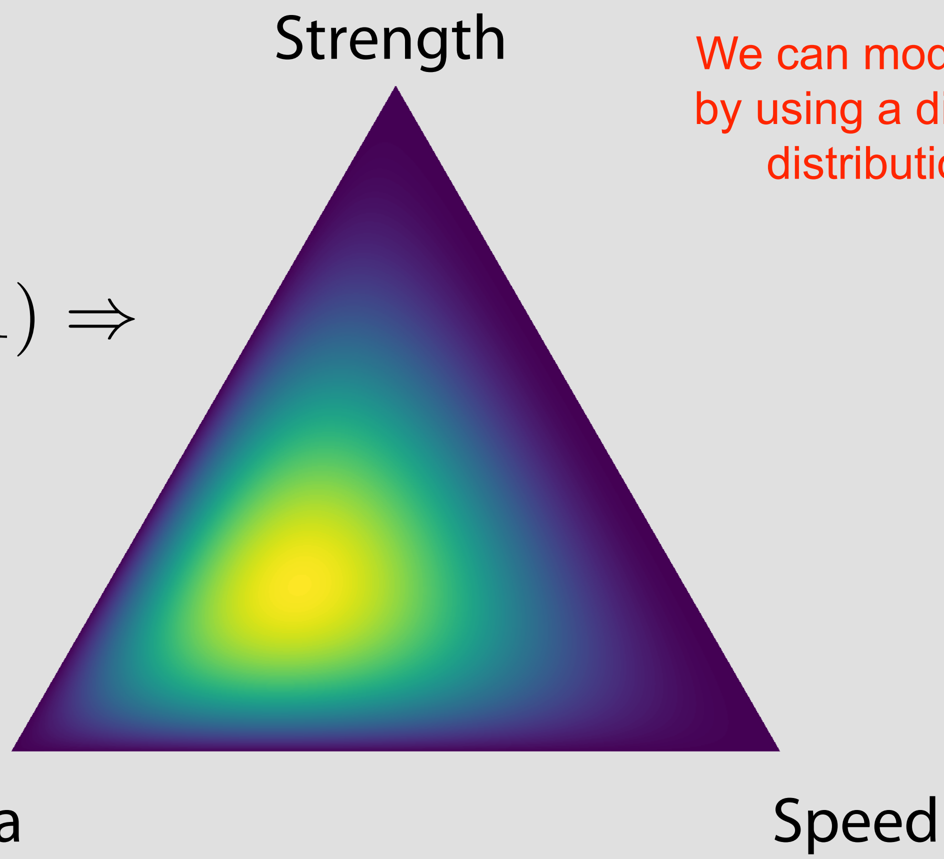


Example

Massively multiplayer online role-playing game (**MMORPG**)

Average over all players:

$$\alpha = (3, 1, 1) \Rightarrow$$



We can model this by using a dirichlet distribution.

Conjugate prior

$P(\theta)$ is **conjugate** to $P(X|\theta)$:

the dirichlet prior is conjugate
to the multinomial likelihood.

$\mathcal{A}(v)$

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

$\mathcal{A}(v)$

What is a multinomial
likelihood? See next slide.

Multinomial likelihood

It is important to consider the combinations in likelihood. In the dice example, each 'side' order doesn't matter. Same like simplices. Each triangle edge likelihood specifically doesn't matter.

Distribution over counts.
e.g. for a dice, K = 6,
each x_i is the number of times that we get number 1.
We conduct n trials.

$$P(X|\theta) = \frac{n!}{x_1! \ldots x_K!} \theta_1^{x_1} \ldots \theta_K^{x_K}$$

sum_i x_i = n.

^ likelihood.

$$p(\theta) = \mathrm{Dir}(\theta|\alpha) = \frac{1}{B(\alpha)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}$$

$$p(\theta|X) \propto \prod_{k=1}^{K} \theta_k^{\alpha_k + x_k - 1}$$

= product of the above.

$$p(\theta|X) = \mathrm{Dir}\left(\theta \middle| \begin{pmatrix} \ldots \\ \alpha_k + x_k \\ \ldots \end{pmatrix}\right)$$