From Gaffke:

ie $Pr(m(X_1 \dots X_n) > 1)$? ($X_1 \dots X_n$ are data).

$$K(x) = P\Big(\sum_{i=1}^{n} x_i D_i \le 1\Big), \tag{9}$$

$D_1, \dots, D_n \sim Dir(\underline{1})$.

where $D = (D_1, \dots, D_n)$ denotes a random variable which is uniformly distributed over the unit simplex in $\mathbb{R}^n$,

$$\Big\{ d = (d_1, \dots, d_n) \in \mathbb{R}^n : d_i \ge 0 \ (1 \le i \le n), \ \sum_{i=1}^{n} d_i \le 1 \Big\};$$

From Rubin's Bayesian Bootstrap:

4. Theory. Let $d = (d_1, \cdots, d_K)$ be the vector of all possible distinct values of X, and let $\theta = (\theta_1, \cdots, \theta_K)$ be the associated vector of probabilities

$$P(X = d_k \mid \theta) = \theta_k, \qquad \sum \theta_k = 1. \tag{4.1}$$

Let $x_1, \cdots, x_n$ be an i.i.d. sample from (4.1) and let n_k = the number of x_i equal to d_k. If the

probability, just as they have zero probability under the sample cdf. The posterior probability for each of the n x_i is centered at $1/n$ but has variability. Specifically, one BB replication is generated by drawing $(n-1)$ uniform $(0, 1)$ random variates $u_1, \cdots, u_{n-1}$, ordering them, and calculating the gaps $g_i = u_{(i)} - u_{(i-1)}$, $i = 1, \cdots, n-1$ where $u_{(0)} = 0$ and $u_{(n)} = 1$. Then $g = (g_1, \cdots, g_n)$ is the vector of probabilities to attach to the data values $x_1, \cdots, x_n$ in that BB replication. Considering all BB replications gives the BB distribution of the distribution of X and thus of any parameter of this distribution.

$g_1, \dots, g_n \sim Dir(\underline{1})$.

For example, with ϕ = mean of X, in each BB replication we calculate the mean of X as if g_i were the probability that $X = x_i$; that is, we calculate $\sum_1^n g_i x_i$. The distribution of the values of $\sum_1^n g_i x_i$ over all BB replications (i.e., generated by repeated draws of the g_i) is the BB distribution of the mean of X.

same as gaffke $m(x_1 \dots x_n)$?

In "Gaffke Bound" notes,

$$m(X_1 \dots X_n) = \sum_{i=1}^{n} X_i D_i + D_{n+1} \quad (X_{n+1} = 1), \quad D_1, \dots, D_{n+1} \sim Dir(\underline{1}).$$

transform to Gaffke

$$m(y_1 \dots y_n) = \sum_{i=1}^{n} y_i h_i, \qquad h_i = \frac{z_i}{\sum_{j=1}^{n+1} z_j}$$

$h_1, \dots, h_n$ is not Dirichlet.

(If wts were $\tilde{h}_i = \dfrac{h_i}{\sum_j h_j}$, it would be $Dir(\underline{1})$)

I'm confused about this.