

18a

Now assume we have

$$G \sim DP(\alpha, H) = DP_{\alpha H}$$

αH is base measure $H(A) = 1$ is prob. dist.

$$\text{we know } G | X_1 \dots X_n \sim DP(\alpha + n, \frac{\alpha H + \sum_i S X_i}{\alpha + n})$$

from earlier arguments.

Another representation:

$$\text{consider } G | X_1 \dots X_n = P$$

~~$P(X_i)$~~ there are k unique X_i , $k < n$

call these $X_1^* \dots X_k^*$

then $P(X_1^*), P(X_2^*), \dots, P(X_k^*), P(A - \sum_i X_i^*)$

$$\sim \text{Dir}(n_1, \dots, n_k, \left[\frac{\alpha H(A - \sum_i X_i^*)}{\alpha + n} \right]_{\alpha + n})$$

$$\approx \text{Dir}(n_1, \dots, n_k, \alpha)$$

This means we can use

represent

as $G(x_1, \dots, x_n)$

$$\sum_{i=1}^k p_i \delta_{x_i}^* + p_{k+1} G^*$$

where $(p_1, \dots, p_{k+1}) \sim \text{Dir}(n_1, \dots, n_k, \alpha)$

and $G_i \sim \text{DP}(\alpha, H)$.