Augment-and-Conquer
Negative Binomial Processes

Mingyuan Zhou and Lawrence Carin

Department of Electrical & Computer Engineering
Duke University, Durham, NC 27708, USA

NIPS, Lake Tahoe, December 05, 2012
We view mixture modeling as a count-modeling problem
- Number of words assigned to topic $k$ in document $j$ (we are modeling a $K \times J$ count matrix in a mixed-membership model)
- The NB distribution $x \sim \text{NB}(r, p)$ can be augmented as
  - a gamma-Poisson mixture distribution
  - a compound Poisson distribution
- We discover a Poisson-logarithmic bivariate count distribution

The joint distribution of the customer count and table count are equivalent:

```
Draw NegBino($r, p$) customers
Assign customers to tables using a Chinese restaurant process with concentration parameter $r$
Draw Poisson($-r \ln (1 - p)$) tables
Draw Logarithmic($p$) customers on each table
```
The NB process $X \sim \text{NBP}(G_0, p)$ augmented as

$$X \sim \sum_{t=1}^{L} \text{Log}(p), \ L \sim \text{PP}(-G_0 \ln(1 - p))$$

is equivalent in distribution to

$$L \sim \text{CRTP}(X, G_0), \ X \sim \text{NBP}(G_0, p).$$

Gamma-NB Process: $X_j \sim \text{NBP}(G, p_j), \ G \sim \text{GaP}(c, G_0)$

Beta-NB Process: $X_j \sim \text{NBP}(r_j, B), \ B \sim \text{BP}(c, B_0)$

Marked-BNBP: $X_j \sim \text{NBP}(R, B), \ (R, B) \sim \text{MBP}(c, G_0, B_0)$

Table: The negative binomial process family: $X_j(\omega_k) \sim \text{NB}(r_{(j,k)}, p_{(j,k)})$.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>$r_k$</th>
<th>$r_j$</th>
<th>$p_k$</th>
<th>$p_j$</th>
<th>$\pi_k$</th>
<th>Related Algorithms</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB-LDA</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td>NMF, LDA, Dir-PFA</td>
</tr>
<tr>
<td>NB-HDP</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>0.5</td>
<td>HDP, DILN-HDP</td>
</tr>
<tr>
<td>NB-FTM</td>
<td>✓</td>
<td></td>
<td></td>
<td>0.5</td>
<td>✓</td>
<td>FTM, S(\gamma)-PFA</td>
</tr>
<tr>
<td>Beta-NB</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>Beta-Geometric, BNBP</td>
</tr>
<tr>
<td>Gamma-NB</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>CRF-HDP</td>
</tr>
<tr>
<td>Marked-Beta-NB</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>BNBP</td>
</tr>
</tbody>
</table>
Figure: Distinct sharing mechanisms and model properties are evident between various NB processes, by comparing their inferred parameters.

Figure: Comparison of per-word perplexities between various algorithms.