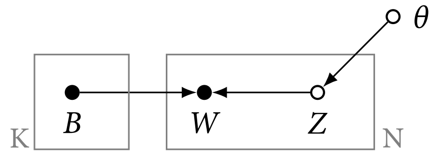
The standard content analysis / topic model each document is modelled as



where W_i is the i-th word of N in the document, Z_i is true topic of W_i, $\theta_k = P(Z = k)$ is the k-th element of length K θ in this document, and β_k is the k-th column of B and is the distribution $P(W \mid Z = k)$.

Let $v(i)$ be the index of the i-th word into the V word vocabulary.

In general

$$
\begin{aligned}
P(Z_i \mid W_i) &= \frac{P(W_i \mid Z_i = k)P(Z_i = k)}{\sum_k^K P(W_i \mid Z_i = k)P(Z_i = k)} \\
&= \frac{\beta_{v(i),k}\theta_k}{\sum_k^K \beta_{v(i),k}\theta_k}
\end{aligned}
$$

However, if topics have exclusive but possibly not exhaustive vocabularies then

$$
\sum_k^K \beta_{v(i),k}\theta_k = \beta_{v(i),k}\theta_k
$$

so if $\theta_k > 0, \ \forall k$ then

$$
P(Z_i \mid W_i) = \mathbb{I}[\beta_{v(i),k} \neq 0]
$$

Measurement error is due to failure of this assumption. Because the estimator of θ_k is

$$
\hat{\theta}_k = \frac{\sum_i^N P(Z_i = k \mid W_i)}{\sum_k^K \sum_i^N P(Z_i = k \mid W_i)}
$$

then each work contributes measurement error associated with each k-generated word is

$$
e_i = 1 - \sum_{j \neq k} P(Z_i = j \mid W_i)
$$

Under measurement error $e_i < 1$, so $\hat{\theta}_k < \theta_k$.

A separate issue from measurement error is bias. At the document level this is

$$
E_k = \hat{\theta}_k - \theta
$$

Note: we can increase error by putting words in the wrong category (measurement error), or by failing to put important words in any category (undercoverage)

Are the precision and recall? Probably