Exercise
Data Science

Bayesian Statistics

Lecturer: Arnim Bleier
arnim.bleier@gesis.org
http://gesis.org/person/arnim.bleier
Bayes’ rule 1

Sum rule: \[ p(C) = \int_M p(C, M) \, dM \]

Product rule: \[ p(C, M) = p(M \mid C) \, p(C) \]

The sum rule and the product rule are the fundamental rules of probability. How can we derive from them Bayes’ rule?

\[
\frac{p(M \mid C)}{p(C)} = \frac{\text{likelihood}}{\text{marginal likelihood}} = \frac{\frac{p(C \mid M)}{p(M)}}{p(C)}
\]
Let's assume a simple world whose inhabitants have one of two mindsets: With a probability of 0.6 they have a ‘liberal’ mindset and with a probability of 0.4 they have a ‘conservative’ one. The liberal minded, in this world, drink latte with a probability of 0.7 and mocha with a probability of 0.3 , and the conservatives drink mocha with a probability of 0.8 and latte with a probability of 0.2 .

We observe someone is drinking mocha. Can we use Bayes’ rule to compute the probability of here/him having a liberal mindset?
Generative Storyline:

\[ \phi_k \sim \text{Dir}(\beta) \quad \forall k \in [1, K] \]

\[ \theta_d \sim \text{Dir}(\alpha) \quad \forall d \in [1, D] \]

\[ z_{di} \sim \text{Cat}(\theta_d) \quad \forall i \in [1, n_d], d \in [1, D] \]

\[ w_{di} \sim \text{Cat}(\phi_{z_{di}}) \quad \forall i \in [1, n_d], d \in [1, D] \]

\[ z_{di} \propto (n_{dk}^{-di} + \alpha) \frac{n_{kw_{di}}^{-di} + \beta}{n_{k.}^{-di} + V \beta} \]
Posterior distribution (LDA)

**Generative Storyline:**

\[ \phi_k \sim \text{Dir}(\beta) \quad \forall k \in [1, K] \]
\[ \theta_d \sim \text{Dir}(\alpha) \quad \forall d \in [1, D] \]

\[ z_{di} \sim \text{Cat}(\theta_d) \quad \forall i \in [1, n_d], d \in [1, D] \]
\[ w_{di} \sim \text{Cat}(\phi_{z_{di}}) \quad \forall i \in [1, n_d], d \in [1, D] \]

What are the parameters of the posterior distribution?