

Coins and Dice Tutorial

Sharad Chitlangia

05-12-2021

Tutorial on Coins and Dice

For a given $F : \mathbb{R} \rightarrow [0, 1]$ is a random variable with a p.d.f $p(F)$, show that $P(heads) = \mathbb{E}[F]$

$$P(heads) = \int_0^1 P(heads|F = f)p(F = f)df$$

$$P(heads) = \int_0^1 fp(f)df$$

$$P(heads) = \mathbb{E}[F]$$

Gamma and Beta Distributions

Gamma

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t} dt$$

$$\Gamma(x+1) = x\Gamma(x)$$

Beta

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$$

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

Beta p.d.f.

$$beta(x; \alpha, \beta) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)}$$

For a random variable $F \sim beta(f; \alpha, \beta)$,

$$\mathbb{E}[F] = \frac{\alpha}{\alpha + \beta}$$

Binomial Sampling

Given a binomial random sample D of size M given a random variable F with p.d.f. $p(F)$ over relative frequencies of heads and a dataset $d = \{x^{(1)}, x^{(2)}, \dots, x^{(M)}\}$ in which there are s values of heads and t values of tails

$$P(d) = \mathbb{E}(F^s [1 - F]^t)$$

If $F \sim p(f) = \text{beta}(f; \alpha, \beta)$, then

$$P(d) = \frac{B(\alpha + s, \beta + t)}{B(\alpha, \beta)}$$

The posterior $P(f|d)$ can also be computed now.

$$P(f|d) = \frac{P(d|f)P(f)}{P(d)}$$

$$P(f|d) = \frac{f^s (1 - f)^t p(f)}{\mathbb{E}(F^s [1 - F]^t)}$$

$$P(f|d) = \text{beta}(f; \alpha + s, \beta + t)$$

$$P(X^{(M+1)} = 1|d) = \frac{\alpha + s}{\alpha + s + \beta + t}$$

Therefore,

$$\text{Prior} = B(\alpha, \beta)$$

$$\text{Likelihood} = \frac{B(\alpha + s, \beta + t)}{B(\alpha, \beta)}$$

$$\text{Posterior} = B(\alpha + s, \beta + t)$$

Relevance to Class Probability Trees (CPT)

We will treat each leaf in a CPT as a coin with its own prior. These coins are independent of coins at other leaves.

Parameter estimation is also straightforward now, if we have the prior and the data. We can directly apply the beta update rule.

For the search procedure, how can we calculate $P(d)$. Recall that calculating this term helps in guiding the search procedure. At each node, we select the successor which increases the value of this term the most (in both trees and bayesian networks).

Conditional Urns Model