Ahmed Bilal

1) $B(a,b) = \int_0^1 \theta^{a-1}(1-\theta)^{b-1}\,d\theta = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

$\Gamma(x+1) = x\Gamma(x)$ ✱ saw from soln

$E(x) = \int_a^b x f_x(x)\,dx = \int_0^1 \left(\frac{1}{B(a,b)}\theta^a (1-\theta)^{b-1}\right) d\theta$

$E(\theta) = \underbrace{\frac{1}{B(a,b)}}_{\text{since this made up of gamma functions which are constants}} \underbrace{\int_0^1 \theta^a (1-\theta)^{b-1}\,d\theta}_{\text{beta function}}$

$= \frac{1}{B(a,b)}\left(B(\alpha+1,\beta)\right) = \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)}$

sub into

$= \left[\frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)}\right]\left[\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right]$

✱ Saw soln $= \left[\frac{a\Gamma(a)\Gamma(b)}{(a+b)\Gamma(a+b)}\right]\left[\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right]$

$= \frac{a}{a+b}$

---

$E[\theta^2] = \frac{1}{B(a,b)}\int_0^1 \theta^{a+1}(1-\theta)^{b-1}\,d\theta = \frac{B(a+2,b)}{B(a,b}$

$= \left[\frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)}\right]\left[\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right]$

☆ Saw solution $= \left[ \frac{a(a+1)\Gamma(a)\Gamma(b)}{(a+b)(a+b+1)\Gamma(a+b)} \right] \left[ \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \right]$

$$= \frac{a(a+1)}{(a+b)(a+b+1)}$$

---

$$Var[\theta] = E[\theta^2] - E[\theta]^2$$

$$= \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2}$$

$$= \frac{a(a+1)(a+b) - a^2(a+b+1)}{(a+b)^2(a+b+1)}$$

$$= \frac{\cancel{a^3} + \cancel{a^2b} + \cancel{a^2} + ab - \cancel{a^3} - \cancel{a^2b} - \cancel{a^2}}{(a+b)^2(a+b+1)}$$

$$= \frac{ab}{(a+b)^2(a+b+1)}$$

---

Mode: $\nabla_\theta P(\theta; a, b) = 0$ ☆ saw soln

$$\nabla_\theta \frac{1}{B(a,b)} = 0$$

$$\nabla_\theta P(\theta; a, b) = \nabla_\theta [\theta^{a-1}(1-\theta)^{b-1}] = 0$$

$$= (a-1)\theta^{a-2}(1-\theta)^{b-1} - \theta^{a-1}(b-1)(1-\theta)^{b-2} = 0$$

☆ saw soln: $(a-1)\theta^{a-2}(1-\theta)^{b-1} = (b-1)\theta^{a-1}(1-\theta)^{b-2}$

$$(a-1)(1-\theta) = (b-1)\theta$$

$a - a\theta - 1 + \theta = b\theta - \theta$ .

$a - 1 = b\theta + a\theta - 2\theta$

$b\theta + a\theta - 2\theta = a - 1$

$(b + a - 2)\theta = a - 1$

$$\theta^* = \frac{a-1}{a+b-2}$$

---

Q 2) $Cat(x|\mu) = \prod_{i=1}^{K} \mu_i^{x_i}$

(step 1: take log)

$\log\left(\prod_{i=1}^{K} \mu_i^{x_i}\right)$

step 2: exponentiate

$Cat(x|\mu) = \exp\left(\log\left(\prod_{i=1}^{K} \mu_i^{x_i}\right)\right)$

$= \exp\left(\sum_{i=1}^{K} \log(\mu_i^{x_i})\right)$

$\exp\left(\sum_{i=1}^{K} x_i(\log \mu_i)\right)$

---

* saw soln: Since $\sum_{i=1}^{K} \mu_i = 1$, and $\sum_{i=1}^{K} x_i = 1$

$\mu_K = 1 - \sum_{i=1}^{K-1} \mu_i$ , $x_K = 1 - \sum_{i=1}^{K-1} x_i$

---

$Cat(x|\mu) = \exp\left(\sum_{i=1}^{K} x_i \log(\mu_i)\right) = \exp\left(\sum_{i=1}^{K-1} x_i \log(\mu_i) + x_K \log(\mu_K\right.$

$= \exp\left[\sum_{i=1}^{K-1} x_i \log(\mu_i) + \left(1 - \sum_{i=1}^{K-1} x_i\right)\log(\mu_K)\right]$

* Saw soln:

$$= \exp\left[\sum_{i=1}^{K-1} x_i \log(\mu_i) - \log(\mu_K) + \log(\mu_K)\right]$$

$$= \exp\left[\sum_{i=1}^{K-1} x_i \log\left(\frac{\mu_i}{\mu_K}\right) + \log(\mu_K)\right]$$

* Saw soln:

let $\eta = \begin{bmatrix} \log\left(\frac{\mu_i}{\mu_K}\right) \\ \vdots \\ \log\left(\frac{\mu_{K-1}}{\mu_K}\right) \end{bmatrix}$

$$\mu_i = \mu_K e^{\eta_i}$$

$$\mu_K = 1 - \sum_{i=1}^{K-1} \mu_i = 1 - \sum_{i=1}^{K-1} \mu_K e^{\eta_i}$$

* Saw soln:

$$= 1 - \mu_K \sum_{i=1}^{K-1} e^{\eta_i} = \frac{1}{1 + \sum_{i=1}^{K-1} e^{\eta_i}}$$

similary,

$$\mu_i = \mu_K e^{\eta_i} = \frac{e^{\eta_i}}{1 + \sum_{i=1}^{K-1} e^{\eta_i}}$$

---

exponential family form: $P(y; \eta) = b(y) \exp(\eta^T T(y) - a(\eta))$

---

$$\mathrm{cat}(x \mid \mu) = (1) \exp(\eta^T x - a(\eta))$$

$$b(\eta) = 1$$

$$T(x) = x$$

Saw soln: $a(\eta) = \log\left(1 + \sum_{i=1}^{K-1} e^{\eta_i}\right)$