## Exercise 2.11

$$\sum_{x=0}^{1} p(x\mid\mu) = \left(\mu^1(1-\mu)^0\right) + \left(\mu^0(1-\mu)^1\right)$$

$$= \mu + (1-\mu) = 1 \quad \textrm{QED}$$

$$\mathbb{E}[x] = \sum_{x=0}^{1} x \cdot \mu^x (1-\mu)^{1-x} = 0 + \mu = \mu$$

$$\textrm{QED}$$

$$VAR[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2 = \sum_{x=0}^{1} x^2 \mu^x (1-\mu)^{1-x} -$$

$$- \mu^2 = 0 + \mu - \mu^2 = \mu(1-\mu)$$

$$\textrm{QED}$$

$$H[x] = -\sum_{x=0}^{1} p(x)\ln(p(x)) =$$

$$= -\sum_{x=0}^{1}\left(\mu^x(1-\mu)^{1-x} \cdot \left(x\cdot\ln(\mu) + (1-x)\ln(1-\mu)\right)\right) =$$

$$= -\left((1-\mu) \cdot \ln(1-\mu) + \mu\cdot\ln(\mu)\right) =$$

$$= -\mu\ln(\mu) - (1-\mu)\ln(1-\mu) \quad \textrm{QED}$$

## Exercise 2.2

$$\sum_{x \in \{-1,1\}} \left(\frac{1-M}{2}\right)^{(1-x)/2} \left(\frac{1+M}{2}\right)^{(x+1)/2} =$$

$$= \left(\frac{1-M}{2}\right) + \left(\frac{1+M}{2}\right) = \frac{1-\cancel{M}+1+\cancel{M}}{2} = 1 \qquad \mathcal{Q}E.D$$

$$\sum_{x \in \{-1,1\}} x \cdot \left(\frac{1-M}{2}\right)^{(1-x)/2} \left(\frac{1+M}{2}\right)^{(x+1)/2} =$$

$$= -\left(\frac{1-M}{2}\right) + \left(\frac{1+M}{2}\right) = \frac{-1+M+x+M}{2} = M \qquad \mathcal{Q}E.D$$

$$Var[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2 = \left(\frac{1-M}{2}\right) + \left(\frac{1+M}{2}\right)$$

$$- M^2 = 1 - M^2$$

$$H[x] = -\sum_{x \in \{-1,1\}} p(x)\ln(p(x)) = -\left(\frac{1-M}{2}\right)\ln\left(\frac{1-M}{2}\right) -$$

$$- \left(\frac{1+M}{2}\right)\ln\left(\frac{1+M}{2}\right)$$

Exercise 2.3

1) show that

(1) $\binom{N}{m} + \binom{N}{m-1} = \binom{N+1}{m} = \dfrac{(N+1)!}{(N+1-m)!\,m!}$

$LHS = \dfrac{N!}{(N-m)!\,m!} + \dfrac{N!}{(N-m+1)!\,(m-1)!} =$

$= \dfrac{N!}{(N-m)!\,m!} + \dfrac{(N+1)!\,m}{(N+1-m)!\,m!\,(N+1)} =$

$= \dfrac{(N+1)!\,(N+1-m)}{(N+1-m)!\,(N+1)\,m!} + \dfrac{(N+1)!\,m}{(N+1-m)\,m!\,(N+1)} =$

$= \dfrac{1}{N+1}\left( \dfrac{(N+1)!\,(N+1-\cancel{m}) + (N+1)!\,\cancel{m}}{(N+1-m)!\,m!} \right) =$

$= \dfrac{1}{\cancel{N+1}}\left[ \dfrac{(N+1)!\cdot N + (N+1)!}{(N+1-m)!\,m!} \right] =$

$= \dfrac{N!\,(N+1)}{(N+1-m)!\,m!} = \dfrac{(N+1)!}{(N+1-m)!\,m!} = \binom{N+1}{m}$

$QED$

2) use this to prove that $^{1)}$

$$(1+x)^N = \sum_{m=0}^{N} \binom{N}{m} x^m$$

· For $N = 0$:

$$(1+x)^0 = \binom{0}{0} x^0 \iff 1 = 1 \cdot 1$$

· Assume that this holds for $N = k$. Now, we show that this assumption $\Rightarrow$ it also holds for $N = k+1$.

$$(1+x)^{k+1} = \sum_{m=0}^{k+1} \binom{k+1}{m} x^m$$

LHS =
$$(1+x)(1+x)^k = \left\{ \begin{array}{c} \text{According to} \\ \text{our assumption} \end{array} \right\} =$$

$$= (1+x) \sum_{m=0}^{k} \binom{k}{m} x^m \quad (1)$$

$$\text{RHS} = \{2.262\} = \sum_{m=0}^{k+1} \left( \binom{k}{m} + \binom{k}{m-1} \right) x^m = \left\{ \begin{array}{c} \binom{k}{k+1} = 0 \\ \binom{k}{-1} = 0 \end{array} \right\} =$$

$$= \sum_{m=0}^{k} \binom{k}{m} x^m + \sum_{m=1}^{k+1} \binom{k}{m-1} x^m = \sum_{m=0}^{k} \binom{k}{m} x^m + x \sum_{n=1}^{k+1} \binom{k}{n-1} x^{n-1}$$

$$(2)$$

(1) $\Rightarrow$ $(1+x) \sum\limits_{m=0}^{K} \binom{k}{m} x^m = \sum\limits_{m=0}^{k} \binom{k}{m} x^m + x \sum\limits_{m=0}^{u} \binom{k}{m} x^m =$

$= \sum\limits_{m=0}^{u} \binom{k}{m} x^m + x \sum\limits_{m=1}^{K+1} \binom{k}{m-1} x^{m-1}$

which is equal
to eq. 2.

QED.

use this to $\checkmark$

Finally, we show that

$\sum\limits_{m=0}^{N} \binom{N}{m} M^m (1-M)^{N-m} = 1$

$LHS = (1-M)^N \sum\limits_{m=0}^{N} \binom{N}{m} M^m (1-M)^{-m} =$

$= (1-M)^N \sum\limits_{m=0}^{N} \binom{N}{m} \left(\frac{M}{1-M}\right)^m = \{2.263\} =$

$= (1-M)^N \left(1 + \frac{M}{1-M}\right)^N = (1-M)^N \left(\frac{1-M+M}{1-M}\right)^N$

$= (1-M)^N \cdot (1-M)^{-N} = (1-M)^0 = 1$   QED.

## Exercise 2.5

1) show that

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

$$\Longleftrightarrow \quad \Gamma(a)\,\Gamma(b) = \Gamma(a+b)\int_0^1 \mu^{a-1}(1-\mu)^{b-1}\, d\mu$$
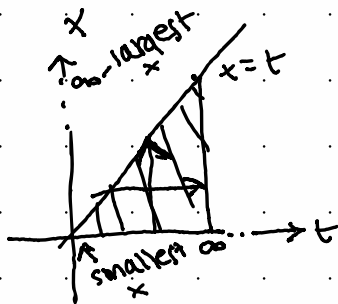
$$LHS = \left[\int_0^\infty \exp(-x)\, x^{a-1}\, dx \int_0^\infty \exp(-y)\, y^{b-1}\, dy\right] =$$

$$= \left[\, t = y+x \,\middle|\, \frac{dt}{dy} = 1 \,\middle|\, \begin{matrix} y=0 \Rightarrow t=x \\ y=\infty \Rightarrow t=\infty \end{matrix}\right]$$

$$= \int_0^\infty \exp(-x)\, x^{a-1} \underbrace{\int_x^\infty \exp(-(t-x))(t-x)^{b-1}\, dt}_{}\, dx$$

cancel

$$= \int_{0=x}^{\infty =x} x^{a-1} \int_{x=t}^{\infty =t} \exp(-t)(t-x)^{b-1}\, dt\, dx = \left[\begin{matrix} \text{change} \\ \text{order} \end{matrix}\right] =$$

$$= \int_0^\infty \int_0^t x^{a-1} \exp(-t)(t-x)^{b-1}\, dx\, dt$$

Finally, we make the change: $x = t\mu$

$$\begin{bmatrix} dx = t \cdot d\mu \\ x \in [0,t] \to \begin{array}{l} \mu = 0 \to x = 0 \\ \mu = 1 \to x = t \end{array} \end{bmatrix}$$

$$\Gamma(a)\,\Gamma(b) = \int_0^\infty \int_0^1 (t\mu)^{a-1} \exp(-t)\,(t - t\mu)^{b-1}\, t\, d\mu\, dt =$$

$$= \int_0^\infty \int_0^1 t^{a-1}\, t^{b-1}\, t\, \exp(-t)\, \mu^{a-1}\, (1-\mu)^{b-1}\, d\mu\, dt =$$

$$= \{ \text{so now we can separate } t \text{ and } \mu \} =$$

$$= \int_0^\infty t^{a+b-1} \exp(-t)\, dt \int_0^1 \mu^{a-1}\, (1-\mu)^{b-1}\, d\mu =$$

$$= \{1.1413 = \Gamma(a+b) \int_0^1 \mu^{a-1}\, (1-\mu)^{b-1}\, d\mu.$$

$$\text{QED.}$$

Note $M$ = avg no heads $\Rightarrow M \in [0,1]$

$$E[M] = \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} M^{a-1} (1-M)^{b-1} \cdot M \, dM =$$

$$= \underbrace{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}}_{\text{const}} \int_0^1 M^{a-1} (1-M)^{b-1} M \, dM =$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 M^{(a+1)-1} (1-M)^{b-1} \, dM =$$

$$= \{2.265\} = \frac{\Gamma(a+b)}{\Gamma(a) \Gamma(b)} \cdot \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+1+b)} =$$

$$= \left\{ \begin{array}{l} \Gamma(c+1) = c \cdot \Gamma(c) \\ \Gamma(c) = (c-1)! \end{array} \right\} = \frac{(a+b-1)! \cdot a \, \Gamma(a)}{\Gamma(a) \, (a+b)!} =$$

$$= \frac{a}{a+b} \qquad \text{QED}$$

$$\text{var}[\mu] = \mathbb{E}[\mu^2] - \underbrace{\mathbb{E}[\mu]^2}_{=\left(\frac{a}{a+b}\right)^2}$$

$$\mathbb{E}[\mu^2] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \underbrace{\mu^{a-1} \cdot \mu^2 \cdot (1-\mu)^{b-1}}_{\mu^{a+2-1}} \, d\mu =$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)} =$$

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

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

$$\text{var}[\mu] = \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2} =$$

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

$$QED$$

## Exercise 2.7

### 1) The posterior mean:

$$\frac{\Gamma(m+a+\ell+b)}{\Gamma(m+a)\,\Gamma(\ell+b)}\int_0^1 \mu^{m+a-1}(1-\mu)^{\ell+b-1}\cdot\mu^1\,d\mu =$$

$$= \left\{2.265\right\} = \frac{\Gamma(m+a+\ell+b)}{\Gamma(m+a)\,\Gamma(\ell+b)}\cdot\frac{\Gamma(m+a+1)\,\Gamma(\ell+b)}{\Gamma(m+a+\ell+b+1)} =$$

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

- The prior mean: $\dfrac{a}{a+b}$

- The maximum livelihood estimate: $\mu_{ML} = \dfrac{m}{N} = \dfrac{m}{m+\ell}$

Now $1$ can be written as:

$$\frac{(m+a)}{(m+a+\ell+b)} = \lambda\cdot\frac{a}{a+b} + (1-\lambda)\frac{m}{m+\ell}$$

⇐⟩

$$\frac{(m+a)}{(m+a+e+b)} = \lambda \cdot \frac{a}{a+b} + (1-\lambda)\frac{m}{m+e} =$$

$$= \lambda \left(\frac{a}{a+b} - \frac{m}{m+e}\right) + \frac{m}{m+e}$$

$$\frac{(m+a)(m+e) - m(m+a+e+b)}{(m+a+e+b)(m+e)} =$$

$$= \lambda \frac{a(m+e) - m(a+b)}{(a+b)(m+e)} \quad \Longleftarrow$$

$$\frac{(\cancel{m^2} + \cancel{me} + am + ae) - \cancel{m^2} - a\cancel{m} - \cancel{me} - mb}{(m+a+e+b)(\cancel{m+e})} =$$

$$= \lambda \frac{a\cancel{m} + ae - \cancel{ma} - mb}{(a+b)(\cancel{m+e})} \quad \Longleftarrow$$

$$\frac{a\,l - mb}{(m+a+e+b)} = \lambda \frac{a\,l - mb}{a+b} \quad \Longleftrightarrow$$

$$\lambda = \frac{a + b}{m + e + a + b}$$

## Exercise 2.8

RHS of 2.270 $= \mathbb{E}_y[\mathbb{E}_x[x|y]] =$

$= \int p(y)\, \mathbb{E}_x[x|y]\, dy = \int\int p(y)\cdot x\, p(x|y)\, dx\, dy =$

$= \int\int x\, p(x,y)\, dx\, dy = \int x\, p(x)\, dx = \mathbb{E}[x]$

QED

Exercise 2.9

- **step 1** : integrate over (2.272):

$$P_{M-1}(\mu_1, \ldots, \mu_{M-2}) = \int_0^{1-\mu_1-\mu_2-\cdots-\mu_{M-2}} P_M(\mu_1, \ldots, \mu_{M-1}) d\mu_{M-1}$$

$$= C_M \int \prod_{k=1}^{M-1}(\mu_k^{\alpha_k-1})\left(1 - \sum_{j=1}^{M-1}\mu_j\right)^{\alpha_M-1} d\mu_{M-1}$$

$$= \left\{\begin{array}{l}\text{All terms that do not depend} \\ \text{on } \mu_{M-1} \text{ are const} \Rightarrow \text{can move outside}\end{array}\right\}$$

$$= C_M\left[\prod_{k=1}^{M-2}(\mu_k^{\alpha_k-1})\right]\int_0^{1-\sum^{M-2}\mu_j}\mu_{M-1}^{\alpha_{k-1}}\left(1 - \sum_{j=1}^{M-1}\mu_j\right)^{\alpha_M-1} d\mu_{M-1}$$

$$\left[\begin{array}{l}\text{Now, we want limits } [0, 1]. \\[4pt] t = \dfrac{\mu_{M-1}}{1-\sum_{j=1}^{M-2}\mu_j} \implies \text{lower} = \dfrac{0}{1-\sum_{j=1}^{M-2}\mu_j} = 0 \\[12pt] \qquad\qquad \text{upper} = \dfrac{1-\sum_{j=1}^{M-1}\mu_j}{1-\sum_{j=1}^{M-2}\mu_j} = 1 \\[12pt] dt = \dfrac{1}{1-\sum_{j=1}^{M-2}\mu_j} d\mu_{M-1}\end{array}\right\} =$$

$$= C_M \left[ \prod_{k=1}^{M-2} M_k^{\alpha_k - 1} \right] \left( 1 - \sum_j^{M-2} M_j \right)^{\alpha_{M-1} - 1} \cdot \left( 1 - \sum_j^{M-2} M_i \right)$$

$$\cdot \int_0^1 t^{\alpha_{M-1} - 1} \left( \underbrace{\left( 1 - \sum_j^{M-2} M_j \right)}_{\text{break out}} - t \cdot \underbrace{\left( 1 - \sum_j^{M-2} M_j \right)}_{\text{break out}} \right)^{\alpha_M - 1} dt =$$

$$= C_M \left\{ \prod_{k=1}^{M-2} M_k^{\alpha_k - 1} \right\} \left( 1 - \sum_j^{M-2} M_i \right)^{\overset{=0}{\overbrace{\alpha_{M-1} - 1 + 1 + \alpha_M - 1}}} \cdot$$

$$\cdot \int_0^1 t^{\alpha_{M-1} - 1} \underbrace{\left( 1 - t \right)^{\alpha_M - 1}}_{\substack{\text{broken} \\ \text{out is}}} dt = \left\{ \begin{array}{c} \text{Exercise} \\ 2.5 \end{array} \right\} =$$

$$= C_M \left[ \prod_{k=1}^{M-2} M_k^{\alpha_k - 1} \right] \left( 1 - \sum_j^{M-2} M_i \right)^{\alpha_{M-1} + \alpha_M - 1} \cdot \frac{\Gamma(\alpha_{M-1}) \Gamma(\alpha_M)}{\Gamma(\alpha_{M-1} + \alpha_M)}$$

so, now we can compare this expression
to (2.38), assuming also that (2.272)
holds for M-1 variables (also see 2.41)

$$C_M \quad \frac{\Gamma(\alpha_{M-1}) \cdot \Gamma(\alpha_M)}{\Gamma(\alpha_{M-1} + \alpha_M)} = \frac{\Gamma(\alpha_1 + \dots + \overbrace{\alpha_{M-1} + \alpha_M})}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_{M-1} + \alpha_M)}$$

$$\Leftrightarrow$$

$$C_M = \frac{\Gamma(\cancel{\alpha_{M-1} + \alpha_M}) \, \Gamma(\alpha_1 + \dots + \alpha_{M-1} + \alpha_M)}{\Gamma(\alpha_1) \bullet \cdots \bullet \Gamma(\cancel{\alpha_{M-1} + \alpha_M}) \, \Gamma(\alpha_{M-1}) \cdot \Gamma(\alpha_M)} =$$

$$= \frac{\Gamma(\alpha_1 + \dots + \alpha_M)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_M)}$$

So, given that the Dirichlet
distribution is normalized for
M-1 variables, it is also normalized
for M variables.

## Exercise 2.10

$$\mathbb{E}[\mu_j] = \int \mu_j \cdot \text{Dir}(\mu; \alpha) \, \underline{d\mu} = \{2.38\}$$

$$= \int \mu_j \cdot \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_k)} \prod_{k=1}^{K} \mu_k^{\alpha_k - 1} \, d\mu =$$

$$= \int \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_u)} \cdot \underbrace{\mu_j^{(\alpha_j - 1) + 1}} \prod_{k \neq j}^{K} \mu_k^{\alpha_u - 1} \, d\mu =$$

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_u)} \cdot \frac{\overbrace{\Gamma(\alpha_1) \cdots \Gamma(\alpha_j + 1) \, \Gamma(\alpha_{j+1}) \cdots \Gamma(\alpha_k)}}{\underset{\downarrow}{\Gamma(\alpha_0 + 1)}}$$

$$\text{since} \, (\alpha_u - 1) + 1 \underline{\underline{=}}$$

$$= \left\{ \text{using} \, \Gamma(x+1) = x \, \Gamma(x) \right\} =$$

$$= \frac{\Gamma(\alpha_0)}{\alpha_0 \Gamma(\alpha_0)} \cdot \frac{\alpha_j \Gamma(\alpha_j)}{\Gamma(\alpha_j)} = \frac{\alpha_j}{\alpha_0} \qquad \text{QED}$$

For the variance:

$$\text{VAR}[\mu_j] = E[\mu_j^2] - \underbrace{E[\mu_j]^2}_{\left(\frac{\alpha_j}{\alpha_0}\right)^2}$$

$$E[\mu_j^2] = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)} \int \mu_j^2 \prod_{k=1}^{K} \mu_k^{\alpha_k - 1} d\mu =$$

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_u)} \int \mu_j^{(\alpha_j + 2) - 1} \prod_{u \neq j}^{K} \mu_u^{\alpha_u - 1} d\mu =$$

$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_u)} \cdot \frac{\Gamma(\alpha_1)\cdots\Gamma(\alpha_j + 2)\cdots\Gamma(\alpha_k)}{\Gamma(\alpha_0 + 2)} =$$

$$= \left\{ \Gamma(x+2) = (x+1) \times \Gamma(x) \right\} =$$

$$= \frac{(\alpha_j + 1)\,\alpha_j}{(\alpha_0 + 1)\,\alpha_0}$$

$$\text{var}[\mu_j] = \frac{\alpha_j(\alpha_j + 1)}{(\alpha_0 + 1)\alpha_0} \;-\; \frac{\alpha_j^2}{\alpha_0^2} = \frac{\alpha_j(\alpha_0 - \alpha_j)}{\alpha_0^2(\alpha_0 + 1)}$$

QED

## Exercise 2.11

$$\cdot \mathbb{E}[\ln \mu_j] = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)} \int \ln \mu_j \prod_{u=1}^{K} \mu_K^{\alpha_u - 1} \, d\vec{\mu} \quad (1)$$

$$\frac{\partial}{\partial j}\left(\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_u)} \prod_{u=1}^{K} \mu_u^{\alpha_u - 1}\right) =$$

$$= \frac{\partial}{\partial j}\left(\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_u)}\right)\cdot \prod_{u=1}^{K} \mu_u^{\alpha_u - 1} \quad +$$

$$+ \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_u)}\cdot \frac{\partial}{\partial j}\left(\prod_{u=1}^{K} \mu_u^{\alpha_u - 1}\right)$$

$$\frac{\partial}{\partial j}\left(\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_u)}\right)\prod_{u=1}^{K}\mu_u^{\alpha_u - 1} \quad +$$

$$+ \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_u)}\cdot \prod_{u\neq j}^{K}\mu_K^{\alpha_u - 1}\cdot \underbrace{\frac{\partial}{\partial j}\left(\mu_j^{\alpha_j - 1}\right)}_{= \mu_j^{\alpha_j - 1}\cdot \ln(\mu_j)}$$

$$(2)$$

Now, we integrate eq. (2):

$$\text{LHS} = \int_{\bar{\mu}} \frac{\partial \, \text{Dir}(\bar{\mu}|\alpha)}{\partial \alpha_j} \, d\bar{\mu} = \frac{\partial}{\partial \alpha_j} \left( \int_{\bar{\mu}} \text{Dir}(\bar{\mu}|\alpha) \, d\mu \right) =$$

$$= \frac{\partial}{\partial \alpha_j} \cdot 1 = 0$$

$$\text{RHS} = \mathbb{E}[\ln \mu_j] + \int \prod_{u=1}^{k} \mu_k^{\alpha_u - 1} \cdot \frac{\partial}{\partial j} \left( \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_k)} \right) d\mu$$

$$\underbrace{\phantom{\frac{\partial}{\partial j} \left( \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_k)} \right)}}_{\substack{\text{const when} \\ \text{integrate wrt} \\ d\mu}}$$

$$= \mathbb{E}[\ln \mu_j] + \frac{\partial}{\partial \alpha_j} \left( \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_k)} \right) \cdot \int_{\bar{\mu}} \prod_{u=1}^{k} \mu_u^{\alpha_u - 1} d\mu$$

$$= \left\{ \begin{array}{l} \text{to keep the dirichlet normalized,} \\ \text{i.e to integrate to one} \Rightarrow \end{array} \right\} =$$

$$= \mathbb{E}[\ln \mu_j] + \frac{\partial}{\partial \alpha_j} \left( \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_u)} \right) \frac{\Gamma(\alpha_1) \cdots \Gamma(\alpha_u)}{\Gamma(\alpha_0)}$$

$$\Longleftrightarrow$$

$$\mathbb{E}[\ln \mu_j] = - \frac{\partial}{\partial \alpha_j} \left( \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_u)} \right) \frac{\Gamma(\alpha_1) \cdots \Gamma(\alpha_u)}{\Gamma(\alpha_0)}$$

$$= \left\{ \frac{\partial}{\partial x} \ln(x) = \frac{1}{x} \right\} = - \frac{\partial}{\partial j} \ln\left( \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)} \right)$$

$$= - \frac{\partial}{\partial \alpha_j} \left[ \ln\left( \Gamma\left( \sum_{k=1}^{K} \alpha_k \right) \right) - \sum_{k=1}^{K} \ln\left( \Gamma(\alpha_k) \right) \right] =$$

$$= \frac{\partial}{\partial \alpha_j} \left( \ln\left( \Gamma(\alpha_j) \right) \right) - \frac{\partial}{\partial \alpha_0} \left( \ln \Gamma(\alpha_0) \right) \cdot \underbrace{\frac{\partial \alpha_0}{\partial \alpha_j}}_{=1} =$$

$$= \psi(\alpha_j) - \psi(\alpha_0) \qquad QED.$$

## Exercise 2.12

1) verify normalized

$$\int_a^b \frac{1}{b-a} \, dx = \left[ \frac{x}{b-a} \right]_a^b = \frac{b-a}{b-a} = 1$$

QED

2) Mean

$$\mathbb{E}[x] = \int_a^b x \cdot \frac{1}{b-a} \, dx = \left[ \frac{x^2}{2} \left( \frac{1}{b-a} \right) \right]_a^b =$$

$$= \frac{1}{(b-a) \cdot 2} \left[ b^2 - a^2 \right] = \frac{(b-a)(b+a)}{(b-a) \cdot 2} =$$

$$= \frac{b+a}{2}$$

## Exercise 2.13]

$$KL(p||q) = -\int N(x|\bar{\mu}, \Sigma) \ln\left(\frac{N(\bar{x}|\bar{\mu}, \Sigma)}{N(\bar{x}|m, L)}\right) dx$$

Let's do this in steps:

$$\ln\left(\frac{N(\bar{x}|\bar{\mu}, \Sigma)}{N(\bar{x}|m, L)}\right) = \frac{1}{2}\ln\frac{|L|}{|\Sigma|} -$$

$$-\frac{1}{2}(\bar{x}-\bar{\mu})^T \Sigma^{-1}(\bar{x}-\mu) + \frac{1}{2}(\bar{x}-\bar{m})^T L^{-1}(\bar{x}-\bar{m})$$

$$KL(p||q) = \frac{1}{2}\ln\frac{|L|}{|\Sigma|} - \frac{1}{2}\int p(x) \cdot (\bar{x}-\bar{\mu})^T \Sigma^{-1}(\bar{x}-\mu)dx +$$

$$+ \frac{1}{2}\int p(x)(\bar{x}-m)^T L^{-1}(\bar{x}-\bar{m}) dx =$$

$$= \int \mathbb{E}[(x-a)^T A(x-a)] = tr(A\Sigma) + (\mu-a)^T A(\mu-a)$$

$$= \frac{1}{2}\ln\frac{|L|}{|\Sigma|} - \frac{1}{2}tr\{I\} + \frac{1}{2}(m-\mu)^T L^{-1}(m-\mu)$$

$$- D$$