

$$\textcircled{1} \quad x \sim \text{Poi}(\mu) ; \quad \mu \sim \text{Ga}(v, 1)$$

use the parametrization  
of T S.1)

a) Before getting the marginal  $f(x)$ ,  
let's get the joint density for  $x, \mu$ .

Joint  $x, \mu$  density

$$f(x, \mu | v) = f(x | \mu, v) \cdot g(\mu | v)$$

$$= f_x(x | \mu) \cdot g(\mu)$$

$$= \frac{\mu^x \cdot e^{-\mu}}{x!} \cdot \frac{\mu^{v-1} \cdot e^{-\mu/1}}{1^v \cdot \Gamma(v)}$$

$$= \frac{\mu^{(v+x)-1}}{\Gamma(x+1)} \cdot e^{-2\mu}$$

(using conditional probability)

Now, get marginal density  $f(x)$  by integrating

over  $\mu$

$$f(x | v) = \frac{1}{\Gamma(x+1)} \cdot \int_0^\infty \mu^{(v+x)-1} \cdot e^{-\mu/(1/2)} d\mu$$

Parameter space for  $\mu$

$$f(x) = \frac{x^{-v-1} - x/\sigma}{\sigma^v \Gamma(v)}$$

$$f(x | v) = \frac{(1/2)^{v+x}}{\Gamma(x+1)} \cdot \int_0^\infty \frac{\mu^{(v+x)-1}}{(1/2)^{v+x}} \cdot e^{-\mu/(1/2)} d\mu$$

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

$$\Gamma(x+2) = x \Gamma(x+1) = x^2 \Gamma(x)$$

→ In our problem ...

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

$$\Gamma(v+2) = v^2 \Gamma(v)$$

: we know  $x$  is an integer  
because it is the

$$\Gamma(v+x) = v^x \Gamma(v)$$

of counts in a Poisson).

$$f(x|v) = \frac{v^x (1/2)^{v+x}}{\Gamma(x+1)} \cdot \int_0^\infty \frac{\mu^{(v+x)-1} e^{-\mu/(1/2)}}{(1/2)^{v+x} v^x \Gamma(v)} d\mu$$

$$= \frac{v^x (1/2)^{v+x}}{\Gamma(x+1)} \cdot \int_0^\infty \frac{\mu^{(v+x)-1} e^{-\mu/(1/2)}}{(1/2)^{v+x} \Gamma(v+x)} d\mu$$

Gamma Density  
 $G_a(v+x, 1/2)$

$$\Rightarrow f(x|v) = \frac{(1/2)^{v+x} \cdot v^x}{\Gamma(x+1)} \cdot \frac{y^a y!}{\Gamma(a+y)}$$

b) Using Bayes Rule

$$f(\mu|x, v) = \frac{f(\mu, x|v)}{f(x|v)}$$

Since we have already computed both  
 $f(\mu, x|v)$  and  $f(x|v)$  it's  
just a matter of simplifying.

(2)

$$\boxed{x \sim F_{v_1, v_2}(x)}$$

$$\Leftrightarrow x \sim \frac{\sqrt{v_2} \text{Ga}(v_1, \sigma)}{\sqrt{v_1} \text{Ga}(v_2, \sigma)}$$

(The F distribution is a renormalized ratio  
of Gamma distributions).

Recall  $\text{Be}(v_1, v_2) \sim \frac{\text{Ga}(v_1, \sigma)}{\text{Ga}(v_1, \sigma) + \text{Ga}(v_2, \sigma)}$

→ Find relationship between  $F_{v_1, v_2}$  and  $\text{Be}(v_1, v_2)$

working with the densities and without loss  
of generality ...

$$f_F(v_1, v_2) = \frac{v_2}{v_1} \cdot \frac{f_{\text{Ga}}(v_1, \sigma)}{f_{\text{Ga}}(v_2, \sigma)}$$

$$\rightarrow \frac{v_1}{v_2} \cdot f_F(v_1, v_2) = \frac{f_{\text{Ga}}(v_1, \sigma)}{f_{\text{Ga}}(v_2, \sigma)}$$

Short handing  $f_{\text{fa}}(v_1, \delta)$  and  $f_{\text{fa}}(v_2, \delta)$   
by  $g_1$  and  $g_2$

Then,  $\frac{v_1}{v_2} \cdot f_F(v_1, v_2) + f_F(v_1, u)$

$$= \frac{g_1}{g_2} + \frac{g_1}{g_1} = \frac{g_1^2 + g_1 g_2}{g_1 g_2} = \frac{g_1(g_1 + g_2)}{g_1 g_2}$$

$$= \frac{(g_1 + g_2)}{g_2}$$

Inverse of  
a Beta distribution

Thus,  $\left[ \frac{v_1}{v_2} \cdot f_F(v_1, v_2) + f_F(v_1, u) \right]^{-1} = \frac{g_2}{g_1 + g_2}$

which is a Beta Distr..

$$x \sim \text{Be}_{v_2, v_1} \iff x \sim \frac{1}{\frac{v_1}{v_2} f_F(v_1, v_2) + f_F(v_1, u)}$$



③ Jupyter

④  $x \sim N_p(\mu, \Sigma)$  (S.14)

→ we start with eq. 5.26

$$\mathbb{E}_\mu = E_\mu \{ \dot{x}_x(\mu) \cdot \dot{x}_x(\mu)^T \}$$

First,

$$f_{\mu}(x) = \frac{(2\pi)^{-p/2}}{|\Sigma|^{1/2}} \cdot \exp\left\{-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu)\right\}$$

pdf of multivariate normal

then

$$\log f_{\mu}(x) = -\frac{p}{2} \log(2\pi) - \frac{1}{2} \log(|\Sigma|)$$

$$-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu)$$

$\ell_x(\mu)$

- We know that  $\ell_x(\mu)$  in our multidimensional context is a gradient. Then  $\ell_x(\mu)$  will be a Hessian matrix (matrix of second derivatives).

→ We'll use that Hessian to more easily compute

$I_{\mu}$

$\ell_x(\mu)$

(Log-likelihood Func)

$\ell_x(\mu)$  =  
(Gradient)

$$I_{\mu} = \begin{bmatrix} \frac{\partial \ell_x(\mu)}{\partial \mu_1} \\ \frac{\partial \ell_x(\mu)}{\partial \mu_2} \\ \vdots \\ \frac{\partial \ell_x(\mu)}{\partial \mu_i} \\ \vdots \\ \frac{\partial \ell_x(\mu)}{\partial \mu_p} \end{bmatrix}$$

$$\ddot{\lambda}_x(\mu) = \begin{bmatrix} \frac{\partial^2 \lambda_x(\mu)}{\partial \mu_1^2} & \frac{\partial^2 \lambda_x(\mu)}{\partial \mu_1 \partial \mu_2} \\ \vdots & \vdots \\ - & - \\ - & - \\ \frac{\partial^2 \lambda_x(\mu)}{\partial \mu_i \partial \mu_j} & \end{bmatrix}$$

Generic element  $\ddot{\lambda}_x(\mu)$  (not diagonal)

Let's calculate generic element  $\ddot{\lambda}_{\mu(ij)}$

Then,  $\frac{\partial^2}{\partial \mu_i \partial \mu_j} \lambda_x(\mu) = \frac{\partial^2}{\partial \mu_i \partial \mu_j} \left( -\frac{1}{2} (x_i - \mu_i) \cdot \sum_{(ij)}^{-1} (x_j - \mu_j) + (x_j - \mu_j) \cdot \sum_{(ji)}^{-1} (x_i - \mu_i) \right)$

$$= \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2, \dots & x_i - \mu_i, \dots \end{bmatrix} \begin{bmatrix} \sum_{(ii)}^{-1} \sum_{(ij)}^{-1} \\ \vdots \\ \sum_{(ci)}^{-1} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 \\ \vdots \\ x_i - \mu_i \end{bmatrix}$$

Constants

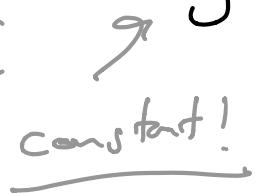
$$(x_i - \mu_i) \cdot \text{constant} \cdot (x_j - \mu_j)$$

Since  $\sum_{(ij)}^{-1} = \sum_{c(j:i)}^{-1}$  (Covariance matrices are symmetrical)

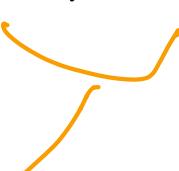
$$\frac{\partial^2}{\partial \mu_i \partial \mu_j} \ell_x(\mu) = \frac{\partial^2}{\partial \mu_i \partial \mu_j} \left[ -(\mathbf{x} : \boldsymbol{\gamma} \boldsymbol{\omega}) \sum_{ij}^{-1} (\mathbf{x}_j \boldsymbol{\gamma} \boldsymbol{\omega}_j)^T \right]$$

$$= -\sum_{(ij)}^{-1}$$

Now, taking  $-E_{\mu} \left[ \frac{\partial^2}{\partial \mu_i \partial \mu_j} \ell_x(\mu) \right] = -E_{\mu} \left[ -\sum_{ij}^{-1} \right]$

Recall, we know the values of  $\sum$   constant!

$$\text{Thus, } I_{\mu(ij)} = -E_{\mu} \left[ -\sum_{ij}^{-1} \right] = \sum_{ij}^{-1}$$

 Generic element of Fisher Inf. matrix

$$\text{and } I_{\mu} = \sum^{-1} \text{  .}$$

Conclusion: Fisher Inf. matrix is the inverse of the known covariance matrix.

- Under repeated sampling, following the same procedure, we get

$$\frac{\partial}{\partial \mu_i \mu_j} l_{\Sigma}(\mu) = \frac{\partial}{\partial \mu_i \mu_j} - \sum_{s=1}^n (x_{is} - \mu_i) \sum_{j=1}^{n-1} (x_{js} - \mu_j)$$

↓  
Sample  $i$ , or  
usual " $i$ ".

$$= -n \sum_{(j)}^{-1}$$

and  $I_{\mu} = n \sum^{-1}$

For large  $n$ , we can connect this result with (5.27) and state

$$\begin{aligned}\hat{\mu}_{\text{anc}} &\sim N(\mu, (n \sum^{-1})) \\ &\sim N(\mu, \Sigma/n)\end{aligned}$$

which is a generalization of the univariate result.

Done.