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

(use the parametrization of T 5.1)

a) Before getting the $\boxed{\text{marginal}}$ of x, let's get the joint density for x, μ.

Jint x, μ density

(using conditional probability)

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

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

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

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

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

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

Parameter space for μ

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

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

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

$\Gamma(x+2) = x\,\Gamma(x+1) = \cancel{x}^2\,\Gamma(x)$

→ In our problem...

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

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

$\vdots$

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

we know x is an <u>integer</u> (because it is the # of counts in a Poisson).

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

$$= \frac{v^x (1/2)^{v+x}}{\Gamma(x+1)} \cdot \int_0^\infty \underbrace{\frac{\mu^{(v+x)-1} e^{-\mu/(1/2)}}{(1/2)^{v+x}\, \Gamma(v+x)}}_{\text{Gamma Density } Ga(v+x,\ 1/2)}\, d\mu$$

$$\Rightarrow f(x|v) = \frac{(1/2)^{v+x} \cdot v^x}{\Gamma(x+1)} \quad \underline{\text{yay!}}$$

b) Using Bayes Rule

$$f(\mu \mid x, \nu) = \frac{f(\mu, x \mid \nu)}{f(x \mid \nu)}$$

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

2

$$\boxed{X \sim F_{\nu_1,\nu_2}(x)}$$

$$\Leftrightarrow \quad X \sim \frac{\nu_2 \, Ga(\nu_1, \sigma)}{\nu_1 \, Ga(\nu_2, \sigma)}$$

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

Recall $\quad Be(\nu_1,\nu_2) \sim \dfrac{Ga(\nu_1,\sigma)}{Ga(\nu_1,\sigma) + Ga(\nu_2,\sigma)}$

→ Find relationship between F_{ν_1,ν_2} and $Be(\nu_1,\nu_2)$ working with the densities and without loss of generality...

$$f_F(\nu_1, \nu_2) = \frac{\nu_2}{\nu_1} \cdot \frac{f_{Ga}(\nu_1, \sigma)}{f_{Ga}(\nu_2, \sigma)}$$

$$\rightarrow \quad \frac{\nu_1}{\nu_2} \cdot f_F(\nu_1, \nu_2) = \frac{f_{Ga}(\nu_1, \sigma)}{f_{Ga}(\nu_2, \sigma)}$$

Shorthanding $f_{Ga}(v_1, \delta)$ and $f_{Ga}(v_2, \delta)$

by g_1 and g_2

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

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

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

Inverse of a Beta distribution

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

which is a Beta Density.

$x \sim Be_{v_2, v_1} \iff x \sim \frac{1}{\frac{v_1}{v_2} f_F(v_1, v_2) + f_F(v_1, v_1)}$

☒

---

③ Jupyter

---

④ $x \sim N_p(\mu, \Sigma)$ (5.14)

-> we start with eq. 5.26

$I_\mu = E_\mu\{\dot{l}_x(\mu) \cdot \dot{l}_x(\mu)^T\}$

First,

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

pdf of MVNormal

then

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

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

- we know that $\dot{\ell}_x(\mu)$ in our multidimensional context is a gradient. Then $\ddot{\ell}_x(\mu)$ will be a Hessian matrix (Matrix of second derivatives).

-> We'll use that Hessian to more easily compute I_μ

$\ell_{\underline{x}}(\mu)$ (Log-likelihood Func)

$$\dot{\ell}_{\underline{x}}(\mu) = \begin{bmatrix} \frac{\partial \ell_x(\mu)}{\partial \mu_1} \\ \frac{\partial \ell_x(\mu)}{\partial \mu_2} \\ \vdots \\ \frac{\partial \ell_x(\mu)}{\partial \mu_i} \\ \vdots \\ \frac{\partial \ell_x(\mu)}{\partial \mu_p} \end{bmatrix}$$

(Gradient)

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

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

Let's calculate generic element $\Sigma_\mu(ij)$

Then, $\frac{\partial^2}{\partial \mu_i \mu_j} l_x(\mu) = \frac{\partial^2}{\partial \mu_i \mu_j} \left\{ -\frac{1}{2} \left[(x_i - \mu_i) \cdot \Sigma^{-1}_{(ij)} (x_j - \mu_j) + (x_j - \mu_j) \cdot \Sigma^{-1}_{(ji)} (x_i - \mu_i) \right] \right\}$

$=$

$$\left[x_1 - \mu_1, x_2 - \mu_2, \dots x_i - \mu_i, \dots\right] \underbrace{\begin{bmatrix} \Sigma^{-1}_{(ii)} & \Sigma^{-1}_{(ij)} \\ & \Sigma^{-1}_{(ij)} \\ & \end{bmatrix}}_{\text{Constants}} \begin{bmatrix} x_1 - \mu_1 \\ \vdots \\ x_i - \mu_i \\ \vdots \end{bmatrix}$$

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

Since $\Sigma^{-1}_{(ij)} = \Sigma^{-1}_{(ji)}$ (covariance matrices are symmetrical)

$$\frac{\partial^2}{\partial \mu_i \mu_j} l_{\underline{x}}(\mu) = \frac{\partial^2}{\partial \mu_i \mu_j}\left[-(x_i - \mu_i)\,\Sigma^{-1}_{ij}\,(x_j - \mu_j)\right]$$

$$= -\Sigma^{-1}_{(ij)}$$

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

Recall, we know the values of Σ $\nearrow$ constant!

Thus, $\underbrace{I_\mu(ij)}_{\text{Generic element of Fisher Inf. matrix}} = -E_\mu\left[-\Sigma^{-1}_{ij}\right] = \Sigma^{-1}_{ij}$

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

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

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

$$\frac{\partial}{\partial\mu_i\partial\mu_j} \ell_{\underline{x}}(\mu) = \frac{\partial}{\partial\mu_i\partial\mu_j} - \sum_{s=1}^{n}(x_{is}-\mu_i)\,\Sigma_{ij}^{-1}\,(x_{js}-\mu_j)$$

Sample #, or usual "i".

$$= -n\,\Sigma^{-1}_{(ij)}$$

and $\boxed{\mathcal{I}_\mu = n\Sigma^{-1}}$

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

$$\hat{\mu}^{MLE} \sim N(\mu, (n\Sigma^{-1})^{-1})$$
$$\sim N(\mu, \Sigma/n)$$

which is a generalization of the univariate result.

<u>Dooone.</u>