Multivariate Normal contined...

Suppose that $X = (x_1, x_2, \dots, x_p)^T$ can be partitioned into

$$X_{(1)} = (X_1, X_2, \dots, X_{p_1})^T$$
$$X_{(2)} = (X_{p_1+1}, X_{p_1+2}, \dots, X_{p_2})^T$$

of course $p = p_1 + p_2$.

Likewise, we can partition μ and Σ. Thus, we can say:

$$\begin{pmatrix} X_{(1)} \\ X_{(2)} \end{pmatrix} \sim N_p\left(\begin{pmatrix} \mu_{(1)} \\ \mu_{(2)} \end{pmatrix}, \begin{pmatrix} \underset{p_1 \times p_1}{\Sigma_{11}} & \underset{p_1 \times p_2}{\Sigma_{12}} \\ \underset{p_2 \times p_1}{\Sigma_{21}} & \underset{p_2 \times p_2}{\Sigma_{22}} \end{pmatrix}\right)$$

If the original vector has a N_p distribution

$$\Rightarrow \underbrace{X_{(2)} \mid X_{(1)}}_{\text{Conditional distribution of } X_{(2)} \text{ given } X_{(1)}} \sim N_{p_2}\Big(\underbrace{\mu_{(2)} + \Sigma_{21}\Sigma_{11}^{-1}(X_{(1)} - \mu_{(1)})}_{\text{mean}},\ \Sigma_{22} - \Sigma_{21}\cdot\Sigma_{11}^{-1}\Sigma_{12}\Big).$$

If $p_1 = p_2 = 1$ (Say I know the height and I want the distribution of the weight)

$$\underset{\text{weight}}{X_2} \mid \underset{\text{Height}}{X_1} \sim N\Big(\mu_2 + \frac{\sigma_{12}}{\sigma_{11}}(X_1 - \mu_1),\ \text{variance}\Big)$$

Mean of weight

Std. dev of height

How much the height deviated from the avg. height

Sqrt of covariance between height and weight

If we let $\mu_1, \mu_2 = 0$ (If we center our data)

$$x_2 \mid x_1 \sim N\left(\underbrace{\frac{\sigma_{12}}{\sigma_{11}}}\, x_1,\ \text{some variance}\right).$$

Linear regression coefficient

$$X_2 = \frac{\sigma_{12}}{\sigma_{11}} \cdot x_1$$

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

$$x_2 \mid x_1 \sim N\left(\text{Mean},\ \sigma_{22} - \frac{\sigma_{21} \cdot \sigma_{12}}{\sigma_{11}}\right)$$

$$\Rightarrow x_2 \mid x_1 \sim N\left(\text{Mean},\ \sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}\right)$$

Note: $\dfrac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}} = r^2$

$$r = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$$

$$\frac{\sigma_{22}^2}{\sigma_{22}} - \frac{\sigma_{12}^2}{\sigma_{11}}$$

$$\frac{\sigma_{22}^2\sigma_{11} - \sigma_{12}^2\sigma_{22}}{\sigma_{22}\sigma_{11}}$$

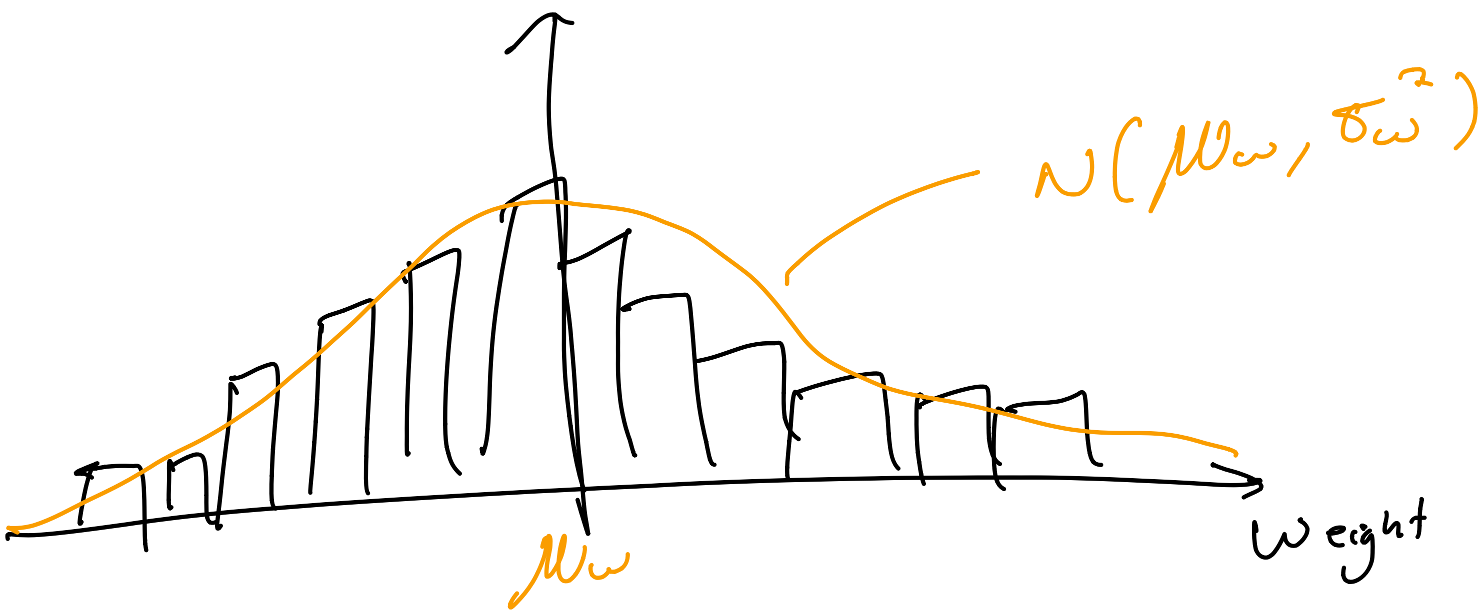
$$\sigma_{22}\left(\frac{\sigma_{22}\sigma_{11}-\sigma_{12}^2}{\sigma_{22}\,\sigma_{11}}\right)$$

$$\sigma_{22}(1-r^2)$$

$$x_2 \mid x_1 \sim N\Big(\text{mean},\ \underbrace{\sigma_{22}}_{\text{std. dev of } x_2}\ \underbrace{(1-r^2)}_{\text{Fraction of unexplained variability.}}\Big)$$

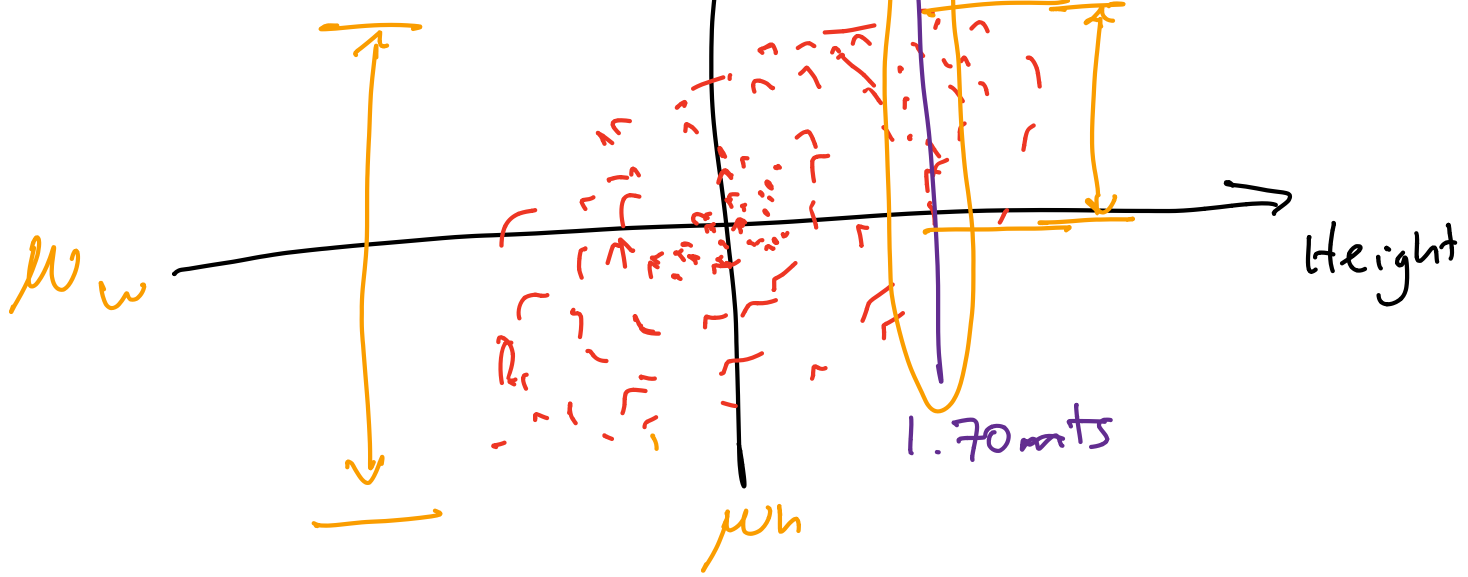
Compare with

$$x_2 \sim N(\text{mean},\ \sigma_{22})$$



$\Downarrow$

If (Height, weight) is MVN



<u>Use of MVN in Bayesian Inference.</u>

First, Univariate case ...

$$f_x(\mu) \sim N(\mu, \sigma^2)$$

$$\mu \sim N(\underbrace{M}_{\text{Prior mean}}, \underbrace{A}_{\text{Prior Variance}})$$

$\Rightarrow$ Can be shown that

$$\underbrace{\mu \mid x}_{\text{Posterior distr for the mean}} \sim N\left(M + \frac{A}{A+\sigma^2}(x-M), \underbrace{\frac{A\sigma^2}{A+\sigma^2}}_{\text{variance}}\right)$$

Expectation of posterior $\left[\hat{\mu}_{Bayes} = M + \frac{A}{A+\sigma^2}(x-M) \right)$

distr.

Multivariate Equivalent is:

$$\mu \sim N_p(M, A) \quad \text{and} \quad f_x(\mu) \sim N_p(\mu, \Sigma)$$

$$\Rightarrow \mu \mid X \sim N_p\left(M + A(A+\Sigma)^{-1}(X-M),\ A(A+\Sigma)^{-1}\Sigma\right)$$

Fisher Information for Multiparameter Families

- The MLE definitions and results for multiparameter case are direct analogues of single parameter results.

Gradient

→ Score Function

$$\dot{l}_x(\mu) = \nabla_\mu \{\log f_\mu(x)\} = \nabla_\mu \{ l_x(\mu) \}$$

$$= \left[\frac{\partial l_x(\mu)}{\partial \mu_1}, \dots, \frac{\partial l_x(\mu)}{\partial \mu_i}, \dots, \frac{\partial l_x(\mu)}{\partial \mu_p} \right]^T.$$

→ Further, it still has mean zero.

$$E_\mu[\dot{l}_x(\mu)] = \underline{0} = [0, 0 \dots, 0]^T$$

- Fisher Information

$$I_\mu = -E_\mu[\ddot{l}_x(\mu)] \checkmark$$

$$\underline{I_\mu} = E_\mu[\dot{l}_x(\mu) \, \dot{l}_x(\mu)^T] = \left(E_\mu \left\{ \frac{\partial l_x(\mu)}{\partial \mu_i} \cdot \frac{\partial l_x(\mu)}{\partial \mu_j} \right\} \right)$$

Hint: $\int_x f(x)[x - \cancel{E(x)}]^2 dx$ (zero)

univariate

$$\downarrow$$

$$\int f(x) \, x^2 \, dx \rightarrow \int_x f(x) [\dot{l}(x)]^2 dx$$

- Fisher Information is a <u>matrix</u> in the multidimensional setting.

- Fisher's Fundamental Theorem of MLE for multidimensional case states:

$$\hat{\mu}^{MLE} \overset{\text{aprox}}{\sim} N_p\left(\mu, I_{\mu}^{-1}\right)$$

as $n \to \infty$ $\blacksquare$