Poission Distribution:

$$f_n(x|\theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \underbrace{\left(\prod_{i=1}^{n} \frac{1}{x_i!}\right)}_{h(x)} \underbrace{e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i}}_{g(t|\theta)}$$

$$MLE = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$\sum_{i=1}^{n} x_i$ is a sufficient statistic for θ.

$$f(x|\theta) = \begin{cases} \theta x^{\theta-1}, & 0<x<1 \\ 0, & \text{Otherwise.} \end{cases}$$

$$f_n(x|\theta) = \underbrace{\theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}}_{g(\theta|x)}$$

$$h(x) = 1$$

$\prod_{i=1}^{n} x_i$ is a sufficient statistic for θ.

Normal distribution, μ is unknown, σ^2 is known.

$$f_n(x|\mu) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}\sigma} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$= \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left[-\frac{\sum(x_i^2 - 2x_i\mu + \mu^2)}{2\sigma^2}\right]$$

$$= \underbrace{\frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{\sum x_i^2}{2\sigma^2}\right)}_{h(x)}$$

$$\underbrace{\frac{\mu}{\sigma^2} \exp\left(\sum_{i=1}^{n} x_i\right) \cdot \exp\left(-\frac{n\mu^2}{2\sigma^2}\right)}_{g(T|x)}$$

$\sum_{i=1}^{n} x_i$ is a sufficient statistics for μ.

or, $\sum_{i=1}^{n} x_i = n\bar{x}$, final expression depends on x only. through the value of $\bar{x}$, therefore $\bar{x}$ is also a sufficient statistics for μ.

Uniform distribution on the interval $[0, \theta]$

$$f(x|\theta) = \begin{cases} \frac{1}{\theta}, & 0 \leq x \leq \theta \\ 0, & \text{otherwise} \end{cases}$$

$$f_n(x|\theta) = \begin{cases} \frac{1}{\theta^n}, & 0 \leq x_i \leq \theta \\ 0, & \text{otherwise} \end{cases}$$

$\mathbb{1}[A]$ is the indicator function of A.

$$f_n(x|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} \mathbb{1}[0 \leq x_i \leq \theta]$$

$$= \frac{1}{\theta^n} \mathbb{1}[\max_i x_i \leq \theta] \, \mathbb{1}[\min_i x_i \geq 0]$$

Then $T(x) = \max_i X_i$ is sufficient statistics for θ.

** here $h(x) = 1$, $g(t|\theta) = I_{[0,\theta]}(t) / \theta^n$

Bernoulé distribution:

$$f_p(x) = p^{\sum_{i=1}^{n} x_i} (1-p)^{n - \sum_{i=1}^{n} x_i}$$

This depends on the data only through $T(x) = \sum_{i=1}^{n} x_i$

$T(X) = \sum_{i=1}^{n} x_i$ is sufficient statistics.

** here $h(x) = 1$

$MLE = \frac{1}{n} \sum_{i=1}^{n} x_i$

Binomial distribution:

$$f(x_i | n, \theta) = {}^{n}C_{x_i} \, \theta^{x_i} (1-\theta)^{n-x_i}$$

$$f(x | n, \theta) = \prod_{i=1}^{n} {}^{n}C_{x_i} \, \theta^{x_i} (1-\theta)^{n-x_i}$$

$$= {}^{n}C_{\sum_{i=1}^{n} x_i} \, \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n^2 - \sum_{i=1}^{n} x_i}$$

$\sum_{i=1}^{n} x_i$ or $\overline{X}$ is the sufficient statistics.

$MLE = \frac{1}{n} \sum_{i=1}^{n} x_i$

Uniform distribution interval $[\alpha, \beta]$

$$f_{X^n}(x^n) = \prod_{i=1}^{n} \left(\frac{1}{\beta-\alpha}\right) 1_{\alpha \leq x_i \leq \beta}$$

$$= \left(\frac{1}{\beta-\alpha}\right)^n 1_{\{\alpha \leq x_i \leq \beta\}}$$

$$= \left(\frac{1}{\beta-\alpha}\right)^n 1_{\{\alpha \leq \min_i x_i\}} 1_{\{\max_i x_i \leq \beta\}}$$

Here $h(x) = 1$.

$$g(t|x) = \left(\frac{1}{\beta-\alpha}\right)^n 1_{\{\alpha \leq \min_i x_i\}} 1_{\{\max_i x_i \leq \beta\}}$$

$h(x)$ doesnot depend on the parameter $(\alpha, \beta\}$, and $g(t|x)$ depends on only x^n through the function $T(x^n) = (\min_i x_i, \max_i x_i)$

Exponential Distribution:

$$f_{X^n}(x^n) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}}$$

$$= \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$

$h(x) = 1$

$$g(t|x) = \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$

$$T(x) = \sum_{i=1}^{n} x_i$$

$$\text{MLE} = \frac{n}{\sum_{i=1}^{n} x_i}$$

Or, $nT(x) = n \cdot \frac{1}{n} \sum_{i=1}^{n} x_i$

$= n\bar{x}$

$\bar{x}$ is a sufficient statistics

Suppose $X_1, X_2, \ldots X_n$ are iid gamma(α, β).

$$f_X(x|\theta) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-\frac{x_i}{\beta}} I(x_i > 0)$$

$$= \underbrace{\left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\right]^n \left(\prod_{i=1}^{n} x_i\right)^\alpha e^{\frac{-\sum_{i=1}^{n} x_i}{\beta}}}_{= g(t_1, t_2|\theta)} \underbrace{\prod_{i=1}^{n} \frac{I(x_i>0)}{x_i}}_{= h(x)}$$

here, $t_1 = \prod_{i=1}^{n} x_i$

$t_2 = \sum_{i=1}^{n} x_i$

$$T = T(x) = \begin{pmatrix} \prod_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i \end{pmatrix} \quad (\alpha, \beta)$$

is sufficient

Suppose $X_1, \ldots X_n$ are iid $U[\theta, \theta+1]$, where $-\infty < \theta < \infty$.

$$f_X(x|\theta) = \prod_{i=0}^{n} I(\theta < x_i < \theta+1)$$

$$= \prod_{i=1}^{n} I(x_i > \theta) \prod_{i=1}^{n} I(x_i - 1 < \theta)$$

$$= \underbrace{I(x_{(1)} > \theta)\, I(x_{(n)} - 1 < \theta)}_{= g(t_1, t_2 | \theta)} \underbrace{\prod_{i=1}^{n} I(x_i \in \mathbb{R})}_{= h(x)}$$

here, $t_1 = X_{(1)} = \min(X_1, X_2, \ldots, X_n)$

$t_2 = X_{(n)} = \max(X_1, X_2, \ldots X_n)$

$$T(X) = \begin{pmatrix} X_{(1)} \\ X_{(n)} \end{pmatrix}$$

Linear regression model $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ for $i = 1, 2, \ldots, n$ where $\epsilon_i \sim$ iid $N(0, \sigma^2)$ and x_i's are fixed constant (not random).

In this model,

$$Y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$$

so, $\theta = (\beta_0, \beta_1, \sigma^2)$

$$f_Y(y|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y_i - \beta_0 - \beta_1 x_i)^2\right\}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2\right\}$$

$$\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)^2=\sum_{i=1}^{n}y_i^2+n\beta_0^2+\beta_1^2\sum_{i=1}^{n}x_i^2$$

$$\underbrace{-2\beta_0\sum_{i=1}^{n}y_i-2\beta_1\sum_{i=1}^{n}x_iy_i+2\beta_0\beta_1\sum_{i=1}^{n}x_i}_{=g(t_1,t_2,t_3\mid\theta)}$$

Here, $t_1=\sum_{i=1}^{n}y_i^2$

$$t_2=\sum_{i=1}^{n}y_i$$

$$t_3=\sum_{i=1}^{n}x_iy_i$$

$$h(y)=1$$

$$T=T(Y)=\begin{cases}\sum_{i=1}^{n}Y_i^2\\ \sum_{i=1}^{n}Y_i\\ \sum_{i=1}^{n}x_iY_i\end{cases}$$

$$f(x\mid\theta)=\begin{cases}\left(\frac{2x}{\theta}\right)e^{-\frac{x^2}{\theta}}, & 0\leq x\leq\infty\\ 0, & \text{elsewhere}\end{cases}$$

$$f(x_1,\dots,x_n;\theta)=\prod_{i=1}^{n}\frac{2x_i}{\theta}e^{-\frac{x_i^2}{\theta}}$$

$$=\frac{2\sum_{i=1}^{n}x_i}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i^2}$$

$$=\underbrace{2\sum_{i=1}^{n}x_i}_{h(x)}\underbrace{\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i^2}}_{g(\theta\mid x)}$$

here, $T(x) = \sum_{i=1}^{n} x_i^2$ sufficient statistics.

$$f(x) = \begin{cases} \theta(\theta+1)\, x^{\theta-1}(1-x), & \text{for } 0<x<1 \\ 0 & , \text{ otherwise.} \end{cases}$$

where, $0<\theta<\infty$

$$f(x|\theta) = \prod_{i=1}^{n} \theta(\theta+1)\, x_i^{\theta-1}(1-x_i)$$

$$= \underbrace{\theta^n(\theta+1)^n \sum_{i=1}^{n} x_i^{\theta-1}}_{g(\theta|x)} \underbrace{\sum_{i=1}^{n}(1-x_i)}_{h(x)}$$

$$= \theta^n(\theta+1)^n \left(\sum_{i=1}^{n} x_i\right)^{\theta-1} \sum_{i=1}^{n}(1-x_i)$$

$$T(x) = \sum_{i=1}^{n} x_i$$

$$f(x) = \begin{cases} 3\theta x^2 e^{-\theta x^3}, & \text{for } 0<x<\infty \\ 0 & , \text{ otherwise} \end{cases}$$

$$f(x;\theta) = \prod_{i=1}^{n} 3\theta\, x_i^2 e^{-\theta x_i^3}$$

$$= 3\,\theta^n \sum_{i=1}^{n} x_i^2\, e^{-\theta \sum_{i=1}^{n} x_i^3}$$

$$= \underbrace{3\,\theta^n e^{-\theta \sum_{i=1}^{n} x_i^3}}_{g(\theta|x)} \underbrace{\sum_{i=1}^{n} x_i^2}_{h(x)}$$

$$T(x) = \sum_{i=1}^{n} x_i^3 \quad \text{sufficient statistic.}$$

$N(0, \sigma^2)$

$$f(x;\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\sum_{i=1}^{n} x_i^2 / 2\sigma^2} \quad [\because \mu = 0]$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi}}\right)^n}_{h(x)} \underbrace{\left(\frac{1}{\sigma}\right)^n e^{-\sum_{i=1}^{n} x_i^2 / 2\sigma^2}}_{g(\sigma^2|x)}$$

$$T(x) = \sum_{i=1}^{n} x_i^2 \quad \text{sufficient statistics.}$$

$$\operatorname{Log} L(x;\sigma^2) = -\frac{n}{2}\operatorname{Log}(2\pi\sigma^2) - \frac{\sum_{i=1}^{n} x_i^2}{2\sigma^2}$$

$$\frac{\partial L(x;\sigma^2)}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{2\cancel{\pi}\sigma^2}\cdot 2\cancel{\pi} + \frac{\sum_{i=1}^{n} x_i^2}{2(\sigma^2)^2} \overset{\text{set}}{=} 0$$

$$\Rightarrow -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^{n} x_i^2}{2\sigma^4} = 0$$

$$\Rightarrow \frac{-n\sigma^2 + \sum_{i=1}^{n} x_i^2}{2\sigma^4} = 0$$

$$\Rightarrow \quad \hat{\sigma}^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$= \frac{T(x)}{n}$$

So, MLE for σ^2 is a function of the sufficient statistic.

Normal Distribution (μ, σ^2)

$$f(x_1, \cdots x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x_i - \mu)^2}{2\sigma^2}$$

$$= \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n \exp - \frac{(\sum x_i - \mu)^2}{2\sigma^2}$$

$$= \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n \exp - \frac{(\sum x_i^2 - 2\sum x_i \cdot \mu + \mu^2)}{2\sigma^2}$$

$$= \left[\frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp - \frac{\mu^2}{2\sigma^2}\right]^n$$

$$\exp\left(-\frac{1}{2\sigma^2}\sum x_i^2 + \frac{\mu}{\sigma^2}\sum x_i\right)$$

$\theta = (\mu, \sigma^2)$ and $h(x) = 1$.

$T(x) = \left(\sum_{i=1}^{n} x_i^2, \sum_{i=1}^{n} x_i\right)$ is a sufficient statistics for (μ, σ^2).

$$\sigma_{MLE} = \frac{1}{n} \sum (x_i - \mu)^2$$

$$\mu_{MLE} = \frac{1}{n} \sum x_i$$

Geometric Distribution:

$$f(X|P) = \prod_{i=1}^{n} (1-p)^{x_i - 1} \, p$$

$$= \underbrace{(1-p)^{\sum x_i - n} \, p^n}_{g(T(x)|p)} \cdot \underbrace{1}_{h(x)}$$

$\sum x_i$ is the sufficient statistics.

$$mle = \frac{n}{\sum x_i}$$

cauchy distribution:

$$f(x, \theta) = \frac{1}{\pi} \cdot \frac{1}{1+(x-\theta)^2} \; ; \; -\infty < x, \theta < \infty$$

$$\prod_{i=1}^{n} f(x_i, \theta) = \left(\frac{1}{\pi}\right)^n \prod_{i=1}^{n} \frac{1}{1+(x_i - \theta)^2}$$

which can't be written in the form $g(T(x)|\theta) h(x)$.

There is no single statistic, which is sufficient estimator for θ.

The order statistics are sufficient for θ.

$$f(x) = \begin{cases} 3\theta x^2 e^{-\theta x^3}, & 0 < x < \alpha \\ 0, & \text{otherwise.} \end{cases}$$

$$f(x;\theta) = \prod_{i=1}^{n} 3\theta x_i^2 e^{-\theta x_i^3}$$

$$= 3\theta^n \sum x_i^2 e^{-\theta^n \sum x_i^3}$$

$$= \underbrace{3\theta^n e^{-\theta^n \sum x_i^3}}_{g(T(x)|\theta)} \underbrace{\sum x_i^2}_{h(x)}$$

$T(x) = \sum x_i^3$ is the sufficient statistics.

$$f(x) = \begin{cases} \frac{\beta}{\theta^3} x^{\beta-1} e^{-\left(\frac{x}{\theta}\right)^\beta} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$f(x_i;\theta) = \prod_{i=1}^{n} \frac{\beta}{\theta^3} x_i^{\beta-1} e^{-\left(\frac{x_i}{\theta}\right)^\beta}$$

$$= \frac{\beta^n}{\theta^{3n}} \sum x_i^{\beta-1} e^{-\sum\left(\frac{x_i}{\theta}\right)^\beta}$$

$$= \left(\frac{\beta}{\theta^3}\right)^n e^{-\sum\left(\frac{x_i}{\theta}\right)^\beta} \underbrace{\sum x_i^{\beta-1}}_{h(x)}$$

$T(x) = \sum x_i^\beta$ is sufficient statistics.

$f(x) = e^{-(x_i - \theta)}$ where $\theta < x_i$

Let $I(X_{(1)}, \theta) = \begin{cases} 1, & \theta < X_{(1)} \\ 0, & \text{otherwise.} \end{cases}$

$$f(x_i|\theta) = e^{-(\Sigma x_i - \theta n)} \cdot I(X_{(1)}, \theta)$$

$$= \underbrace{e^{-\Sigma x_i}}_{h(x)} \cdot e^{\theta n} \cdot I(X_{(1)}, \theta)$$

$X_{(1)}$ is the sufficient statistics.

$e^{(x-\theta)}$; $x \geq \theta$

$$f(x;\theta) = \prod_{i=1}^{n} e^{(x_i - \theta)}$$

$$= e^{\Sigma x_i - \theta n}$$

$$= \underbrace{e^{\Sigma x_i} \cdot e^{-\theta n}}_{g(T(x)|\theta)} \cdot \underbrace{1}_{h(x)}$$

So. Σx_i or $\bar{x}$ is the sufficient statistics.

another way:

$$f(x;\theta) = e^{\sum x_i} e^{-\theta n} \cdot e^{-n} \cdot e^{n}$$

$$= e^{n\bar{x}-n} e^{-\theta n} \cdot e^{n}$$

$$= \underbrace{e^{n(\bar{x}-1)} \cdot e^{-\theta n} \cdot e^{n}}_{g(T(x)|\theta)} \cdot \underset{h(x)}{1}$$

So, $\bar{x}-1$ is the sufficient statistics.

$$f(x,p) = p^{x} q$$

$$f(x, p) = \prod_{i=1}^{n} p^{x_i} q = \underbrace{p^{\sum x_i} q^{n}}_{g(T(x)|p)} \cdot \underbrace{1}_{h(x)}$$

$\sum x_i$ is the sufficient statistics.

Beta distribution:

$$f(\alpha,\beta;x) = \prod_{i=1}^{n} \frac{1}{B(\alpha,\beta)} x_i^{\alpha-1} (1-x_i)^{\beta-1}$$

$$= \left(\frac{\Gamma(a)\Gamma(\beta)}{\Gamma(a+b)}\right)^{n} \left(\sum x_i\right)^{\alpha-1} \left(\sum (1-x_i)\right)^{\beta-1}$$

$h(x) = 1.$

$\{\sum x_i, \sum (1-x_i)\}$ are sufficient statistics for (α, β).

$$f(x) = \frac{1}{\theta^2} x e^{-\frac{x}{\theta}}$$

$$f(x;\theta) = \prod_{i=1}^{n} \frac{1}{\theta^2} x_i e^{\frac{-x_i}{\theta}}$$

$$= \frac{1}{\theta^{2n}} \sum x_i \, e^{-\sum \frac{x_i}{\theta}}$$

$$= \underbrace{\frac{1}{\theta^{2n}} e^{-\frac{1}{\theta} \sum x_i}}_{g(T(x)|\theta)} \cdot \underbrace{\sum x_i}_{h(x)}$$

$T(x) = \sum x_i$ is sufficient statistic.

Raylieg h Distribution:

$$f(x,\sigma^2) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$$

$$\prod_{i=1}^{n} \frac{x_i}{\sigma^2} e^{-x_i^2/2\sigma^2}$$

$$= \frac{\sum x_i}{\sigma^{2n}} e^{-\frac{\sum x_i^2}{2\sigma^2}}$$

$$= \underbrace{\sum x_i}_{h(x)} \underbrace{\frac{1}{\sigma^{2n}} e^{-\frac{1}{2\sigma^2} \sum x_i^2}}_{g(T(x)|\sigma)}$$

$T = \sum x_i^2$ Sufficient statistics.

Pareto Distribution:

$$f(x,\alpha) = \frac{\alpha K^\alpha}{x^{\alpha+1}}.$$

$$\prod_{i=1}^{n} \frac{\alpha K^\alpha}{x_i^{\alpha+1}}$$

$$= \frac{\alpha^n K^{\alpha n}}{\sum x_i^{\alpha+1}} \cdot \underbrace{1}_{h(x)}$$

$$= g(T(x)|\alpha) \quad h(x)$$

$T = \sum x_i$ is sufficient statistics.

Beta Distribution:

$$\frac{1}{\Gamma(\alpha)\theta^{\alpha}} y^{\alpha-1} e^{-y/\theta}$$

$$\prod_{i=1}^{n} \left(\frac{1}{\Gamma(\alpha)\theta^{\alpha}}\right) y_i^{\alpha-1} e^{-y_i/\theta}$$

$$= \left(\frac{1}{\Gamma(\alpha)\theta^{\alpha}}\right)^n \Sigma y_i^{\alpha-1} e^{-\Sigma y_i/\theta}$$

$$= \underbrace{\left(\frac{1}{\Gamma(\alpha)\theta^{\alpha}}\right)^n e^{-\Sigma y_i/\theta}}_{g(T(x)|\theta)} \cdot \underbrace{\Sigma y_i^{\alpha-1}}_{h(x)}$$

$T = \Sigma y_i$ sufficient statistics