Theorem (Fischer-Neyman Factorization)

Let $\underline{X} = (X_1, \dots, X_v)$ r.s of distribution with PDF $f(\underline{x}, \underline{\theta})$
$\underline{\theta} = (\theta_1, \dots, \theta_s)$. Then the statistic $\underline{T}(\underline{x}) = (T_1(\underline{x}), T_2(\underline{x}), \dots, T_k(\underline{x}))$
is sufficient for $\underline{\theta}$ iff $\exists g, h$ such that :
$f(\underline{x} ; \underline{\theta}) = g(T(\underline{x}) ; \underline{\theta}) \cdot h(\underline{x})$
(g depends on x only via $T(x)$ and h is independent of $\underline{\theta}$)

Proof

(We will prove it for a discrete distribution
and $v=1$, $s=1$, $k=1$ for simplicity)

($\Leftarrow$)

We suppose that $f(x ; \theta) = g(T(x) ; \theta) \cdot h(x)$ and will prove that $T(x)$ is sufficient for θ.

$$P(X=x \mid T=t) = \begin{cases} \dfrac{P(X=x, T=t)}{P(T=t)} & , \text{ if } T=T(x)=t \\ 0 & , \text{ otherwise} \end{cases}$$

and we have $P(T=t) = \sum\limits_{T(x)=t} P(X=x) = \sum\limits_{T(x)=t} f(x ; \theta) =$
$= \sum\limits_{T(x)=t} g(T(x) ; \theta) \cdot h(x) = g(t ; \theta) \cdot \sum\limits_{T(x)=t} h(x)$

so for $T(x)=t$: $P(X=x \mid T=t) = \dfrac{P(X=x, T=t)}{P(T=t)} = \dfrac{P(X=x)}{P(T=t)} =$
$= \dfrac{g(T(x), \theta) \cdot h(x)}{g(T(x), \theta) \cdot \sum\limits_{T(x)=t} h(x)} = \dfrac{h(x)}{\sum\limits_{T(x)=t} h(x)}$, which is independent of θ

so $T(x)$ is sufficient for θ.

($\Rightarrow$) We suppose that $T(x)$ is sufficient for θ. Then:

$f(x;\theta) = P(X=x) = P(X=x, T=t) = P(X=x|T=t)\cdot P(T=t)$

Since $T(x)$ is sufficient $P(X=x, T=t)$ is independent of θ. Hence $f(x;\theta) = g(T(x);\theta)\cdot h(x)$, where $h(x) = P(X=x, T=t)$ and $g(T(x);\theta) = P(T=t)$ $\square$

Examples

1) Let $X_1, X_2, \dots, X_n$ r.s of Poisson(λ). Find a sufficient statistic for λ.

Solution

The joint PDF of $X_1, \dots, X_n$ is:

$f(\underline{x};\lambda) = \prod_{i=1}^{n} f(x_i;\lambda) = \prod_{i=1}^{n} e^{-\lambda}\cdot\frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda}\cdot\frac{\lambda^{\sum x_i}}{\prod x_i!} =$

$= g(T(\underline{x});\lambda)\cdot h(\underline{x})$

where $g(T(\underline{x});\lambda) = e^{-n\lambda}\cdot\lambda^{\sum_{i=1}^{n} x_i}$, $h(x) = \left(\prod_{i=1}^{n} x_i!\right)^{-1}$

so $T(\underline{x}) = \sum_{i=1}^{n} X_i$ is a sufficient statistic for λ.

2) Let $X_1, \dots, X_n$ r.s with PDF $f(x;\theta) = \theta\cdot x^{\theta-1}$, $0<x<1$, $\theta>0$

Find a sufficient statistic for θ

Solution

$f(x;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n}\theta x_i^{\theta-1} = \theta^{n}\left(\prod_{i=1}^{n} x_i\right)^{\theta-1} = g(T(\underline{x}),\theta)\cdot h(\underline{x})$

where $h(\underline{x}) = 1$ and $g(T(\underline{x}),\theta) = \theta^{n}\left(\prod_{i=1}^{n} x_i\right)$ so $T(\underline{x}) = \prod_{i=1}^{n} x_i$ is sufficient for θ

3) Let $X_1, \dots, X_n$ r.s of $N(\mu, \sigma^2=1)$. Find a sufficient statistic for μ.

Solution

$$f(\underline{x};\mu)=\prod_{i=1}^{\nu} f(x_i;\mu)=\prod_{i=1}^{\nu}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(x_i-\mu)^2}{2}\right\}=$$

$$=(2\pi)^{-\nu/2}\exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{\nu}x_i^2+\nu\mu^2-2\mu\sum_{i=1}^{\nu}x_i\right]\right\}=$$

$$=(2\pi)^{-\nu/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{\nu}x_i^2\right\}\cdot\exp\left\{-\frac{\nu}{2}\mu^2+\mu\cdot\sum_{i=1}^{\nu}x_i\right\}=$$

$$=\underbrace{(2\pi)^{-\nu/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{\nu}x_i^2\right\}}_{h(\underline{x})}\cdot\underbrace{\exp\left\{-\frac{\nu}{2}\mu^2+\mu\cdot\sum_{i=1}^{\nu}x_i\right\}}_{g\left(T(\underline{x})=\sum_{i=1}^{\nu}x_i,\ \mu\right)}$$

Hence $T(\underline{x})=\sum_{i=1}^{\nu}x_i$ is a sufficient statistic for μ

Alternatively: $f(\underline{x};\mu)=(2\pi)^{-\nu/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{\nu}(x_i-\bar{x}+\bar{x}-\mu)^2\right\}=$

$$=(2\pi)^{-\nu/2}\exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{\nu}(x_i-\bar{x})^2+\sum_{i=1}^{\nu}(\bar{x}-\mu)^2+2(\bar{x}-\mu)\cdot\sum_{i=1}^{\nu}(x_i-\bar{x})\right\}.$$

We have $\sum_{i=1}^{\nu}(x_i-\bar{x})=\sum_{i=1}^{\nu}x_i-\nu\bar{x}=\nu\bar{x}-\nu\bar{x}=0$

so $f(\underline{x};\mu)=(2\pi)^{-\nu/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{\nu}(x_i-\bar{x})^2\right\}\cdot\exp\left\{-\frac{1}{2}\sum_{i=1}^{\nu}(\bar{x}-\mu)^2\right\}=$

$$=h(\underline{x})\cdot g(T(\underline{x})=\bar{x};\mu)$$

hence $T(\underline{X})=\bar{X}$ is sufficient for μ

4) Let $X_1,X_2,\dots,X_\nu$ r.s. of $N(\mu,\sigma^2)$. Find a sufficient statistic for $\underset{\sim}{\theta}=(\theta_1=\mu,\theta_2=\sigma^2)$.

Solution

$$f(\underline{x};\mu,\sigma^2)=(2\pi\sigma^2)^{-\nu/2}\cdot\exp\left\{-\frac{1}{2\sigma^2}\cdot\sum_{i=1}^{\nu}(x_i-\bar{x}-\bar{x}-\mu)^2\right\}=$$

$$=(2\pi\sigma^2)^{-\nu/2}\cdot\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{\nu}(x_i-\bar{x})^2-\frac{1}{2\sigma^2}\sum_{i=1}^{\nu}(\bar{x}-\mu)^2+\frac{1}{\sigma^2}(\bar{x}-\mu)\cancelto{0}{\sum_{i=1}^{\nu}(x_i-\bar{x})}\right\}$$

$$=h(\underline{x})\cdot g\left(T_1(\underline{x})=\bar{x},\ T_2(\underline{x})=\sum_{i=1}^{\nu}(x_i-\bar{x})^2,\ \mu,\sigma^2\right)$$

so $\underline{T}(\underline{x})=\left(\bar{X},\sum_{i=1}^{\nu}(x_i-\bar{x})^2\right)$ is sufficient for $\underset{\sim}{\theta}$

Note

exponential family of distributions

Let $X_1,\dots,X_\nu$ r.s. of EFD. Can we find a sufficient statistic for $\underline{\theta}=(\theta_1,\dots,\theta_s)$?

If $X_j \sim$ EFD with $\underline{\theta}=(\theta_1,\dots,\theta_s)$, $j=1,\dots,\nu$ then
$f_{X_j}(x_j;\underline{\theta}) = B(\underline{\theta})\cdot \exp\{\sum_{i=1}^{s} q_i(\underline{\theta})T_i(x_j)\}\cdot h(x_j)$
The joint distribution of $\underline{X}=(X_1,\dots,X_\nu)$ belongs to the ν-dimesional EFD with s parameters. Hence:
$$f_{\underline{X}}(\underline{x};\underline{\theta}) = \prod_{j=1}^{\nu} f(x_j;\underline{\theta}) = [B(\underline{\theta})]^{\nu}\cdot \exp\{\sum_{j=1}^{\nu}\sum_{i=1}^{s} q_i(\underline{\theta})T_i(x_j)\}\cdot\prod_{j=1}^{\nu} h(x_j) =$$
$$= B^*(\underline{\theta})\cdot \exp\{\sum_{i=1}^{\nu} q_i(\underline{\theta})T_i^*(\underline{x})\}\cdot h^*(\underline{x})$$
where $B^*(\underline{\theta}) = (B(\underline{\theta}))^{\nu}$, $h^*(\underline{x}) = \prod_{j=1}^{\nu} h(x_j)$, $T_i^*(\underline{x}) = \sum_{j=1}^{\nu} T_i(x_j)$

so Neyman's theorem applies and $\underline{T}(\underline{x}) = (T_1^*(\underline{x}),\dots,T_s^*(\underline{x}))$ is sufficient for $\underline{\theta}$.

Corollary

Let $\varphi: \mathbb{R}^k \to \mathbb{R}^k$ and $\omega: \mathbb{R}^s \to \mathbb{R}^s$, where φ,ω "1-1" (so that φ^{-1},ω^{-1} exist). If the statistic $\underline{T}=\underline{T}(\underline{x})$ is sufficient for $\underline{\theta}$ then:
i) $\underline{T}_1 = \varphi(\underline{T})$ is sufficient for $\underline{\theta}$
ii) $\underline{T}$ is sufficient for $\underline{\theta}_1 = \omega(\underline{\theta})$

Proof

ii) $f(\underline{x};\underline{\theta}) = g(\underline{T}(\underline{x}),\underline{\theta})\cdot h(\underline{x}) = g(\underline{T}(\underline{x}),\omega^{-1}(\underline{\theta}_1))\cdot h(\underline{x}) = g_2(\underline{T}(\underline{x}),\underline{\theta}_1)h(\underline{x})$
so $\underline{T}$ is sufficient for $\underline{\theta}_1$

i) $f(\underline{x};\underline{\theta}) = g(\underline{T}(\underline{x}),\underline{\theta})\cdot h(\underline{x}) = g(\varphi^{-1}(\underline{T}_1(\underline{x})),\underline{\theta})\cdot h(\underline{x}) =$
$= g_1(\underline{T}_1(\underline{x}),\underline{\theta})\cdot h(\underline{x})$
so $\underline{T}_1$ is sufficient for $\underline{\theta}$

Example

1) If $T=\sum_{i=1}^{n} x_i$ is sufficient for θ then:

$\rightsquigarrow \frac{T}{n}=\frac{\sum x_i}{n}=\bar{X}$ sufficient for θ

$\rightsquigarrow T$ is sufficient for $\theta^3, 2\theta, X\theta, \log\theta, e^{\theta}$

2) Let $X_1,\dots,X_n$ r.s of $U(0,\theta)$. Find a sufficient statistic for θ.

Solution

$U(0,\theta)$ is not in EFD so we can't work with that.

Generally, if $X\sim U(a,\beta)$ then $f(x;a,\beta)=\frac{1}{\beta-a}$, $a<x<\beta$

or alternatively $f(x;a,\beta)=\frac{1}{\beta-a}\cdot I(a<x<\beta)$

where $I(a<x<\beta)=\begin{cases}1, & a<x<\beta\\ 0, & \text{otherwise}\end{cases}$

For $X\sim U(0,\theta)$ it is $f(x;\theta)=\frac{1}{\theta}I(0<x<\theta)$

So according to Neyman Factorization we have:

$$f(\underline{x}_n;\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\prod_{i=1}^{n}\frac{1}{\theta}I(0<x_i<\theta)=\theta^{-n}\prod_{i=1}^{n}I(0<x_i<\theta)$$

We want $\prod_{i=1}^{n}I(0<x_i<\theta)=1$ or else $f(\underline{x},\theta)=0$

so we want $I(0<x_i<\theta)=1 \quad \forall i=1,\dots,n \Leftrightarrow$

$\Leftrightarrow x_i\in(0,\theta) \quad \forall i \Leftrightarrow 0<X_{(1)}<X_{(2)}<\dots<X_{(n)}<\theta$

where $X_{(1)},X_{(2)},\dots,X_{(n)}$ are the sorted statistical data

and $X_{(1)}=\min\{x_i, i=1,\dots,n\}$, $X_{(n)}=\max\{x_i, i=1,\dots,n\}$

Hence we want $I(0<X_{(1)}<+\infty)\cdot I(0<X_{(n)}<\theta)=1$

so $f(\underline{x};\theta)=\theta^{-n}I(0<X_{(1)}<+\infty)\cdot I(0<X_{(n)}<\theta)=$

$=g(T(\underline{x});\theta)\cdot h(\underline{x})$, where $g(T(\underline{x});\theta)=\theta^{-n}I(0<X_{(n)}<\theta)$

and $h(\underline{x})=I(0<X_{(1)}<+\infty)$

Finally, $T = T(\underline{x}) = X_{(n)}$ is a sufficient statistic for θ

Alternatively, we can demand $I(0 < x_{(1)} < \theta) \cdot I(0 < x_{(n)} < \theta) = 1$ and we would get $f(\underline{x};\theta) = \theta^{-n} I(0 < x_{(1)} < \theta) \cdot I(0 < x_{(n)} < \theta)$ so $h(\underline{x}) = 1$ and $\underline{T_1}(\underline{x}) = (x_{(1)}, x_{(n)})$ which is a sufficient statistic but not the minimal.