Rao-Blackwell and sufficiency

Recap: Rao-Blackwell

Let θ be a parameter of interest, and $\tau(\theta)$ some function of θ. Let $\hat{\tau}$ be some unbiased estimator of $\tau(\theta)$, and T a sufficient statistic for θ.

Let $\tau^* = \mathbb{E}[\hat{\tau} \mid T]$. Then:

1. $\mathbb{E}[\tau^*] = \tau$ (unbiased)
2. $Var(\tau^*) \leq Var(\hat{\tau})$

Reminder: If the conditional distribution of $X_1, \ldots, X_n \mid T$ does not depend on θ, then T is a sufficient statistic for θ

Proof: ① $\mathbb{E}[\tau^*] = \mathbb{E}[\mathbb{E}[\hat{\tau} | T]] = \mathbb{E}[\hat{\tau}] = \tau(\theta)$ ✓

(iterated expectation)

② $$\begin{aligned} Var(\tau^*) &= Var(\mathbb{E}[\hat{\tau} | T]) \\ &\leq Var(\mathbb{E}[\hat{\tau} | T]) + \overbrace{\mathbb{E}[Var(\hat{\tau} | T)]}^{\geq 0} \\ &= Var(\hat{\tau}) \end{aligned}$$ (law of total variance) ✓

why do we need sufficiency? If τ^* is a function of θ, we can't actually calculate τ^*. But if T is a sufficient statistic for θ, then $(X_1, \dots, X_n) | T$ does not depend on θ

$\Rightarrow$ $\hat{\tau} | T$ does not depend on θ

$\Rightarrow$ $\underbrace{\mathbb{E}[\hat{\tau} | T]}_{\tau^*}$ cannot involve θ //

Example: Rao-Blackwell

Let $X_1, \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$.

a sufficient statistic: $T = \sum_i X_i$

Consider $\hat{\lambda} = X_1$ $\quad \mathbb{E}[\hat{\lambda}] = \lambda$ $\quad Var(\hat{\lambda}) = \lambda$

$T = \sum_i X_i$ $\quad \lambda^* = \mathbb{E}[\hat{\lambda} \mid T]$

Now, $\underbrace{\mathbb{E}[\textstyle\sum_i X_i \mid T = t]} = \mathbb{E}[T \mid T = t] = t$

$\Rightarrow \sum_i \mathbb{E}[X_i \mid T = t] = t$

$\Rightarrow \mathbb{E}[X_1 \mid T = t] = \dfrac{t}{n}$

$\Rightarrow \lambda^* = \dfrac{T}{n} = \dfrac{1}{n} \sum_{i=1}^{n} X_i$

Conditioning on an insufficient statistic:

$$X_1, \dots, X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$$

X_2 is <u>not</u> sufficient for λ

$$\hat{\lambda} = X_1$$

$$\lambda^* = \mathbb{E}[X_1 \mid X_2]$$

$$= \mathbb{E}[X_1] = \lambda$$

<u>not</u> a valid estimator of λ

X_1, X_2 $\quad P(X_1=1, X_2=3, X_1+X_2=4)$

$= P(X_1=1, X_2=3)$

Example: sufficiency

Let $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$.

> Find a sufficient statistic for p.

$T = \frac{1}{n}\sum_i X_i$ or $T = \sum_i X_i \sim Binomial(n, p)$

e.g. $T = \sum_i X_i$ wTS: $f(X_1, \ldots, X_n \mid \sum_i X_i, p)$ does not depend on p

$$f(X_1, \ldots, X_n \mid \textstyle\sum_i X_i, p) = \frac{f(X_1, \ldots, X_n, \sum_i X_i \mid p)}{f(\sum_i X_i \mid p)}$$

$$= \frac{f(X_1, \ldots, X_n \mid p)}{f(\sum_i X_i \mid p)} = \frac{\prod_i p^{X_i}(1-p)^{1-X_i}}{\binom{n}{\sum_i X_i} p^{\sum_i X_i}(1-p)^{n-\sum_i X_i}}$$

$$= \frac{\cancel{p^{\sum_i X_i}(1-p)^{n-\sum_i X_i}}}{\binom{n}{\sum_i X_i}\cancel{p^{\sum_i X_i}(1-p)^{n-\sum_i X_i}}} = \frac{1}{\binom{n}{\sum_i X_i}}$$

(does not depend on p)

Factorization theorem

Let $X_1, \dots, X_n$ be a sample with joint probability function $f(x_1, \dots, x_n \mid \theta)$. A statistic $T \equiv T(X_1, \dots, X_n)$ is sufficient for θ if and only if there exist functions $g(t \mid \theta)$ and $h(x_1, \dots, x_n)$ such that, for all possible $x_1, \dots, x_n$ and all possible θ,

$$f(x_1, \dots, x_n \mid \theta) = \underbrace{g(T(x_1, \dots, x_n) \mid \theta)}_{\text{joint distribution only depends on } \theta \text{ through the sufficient statistic}} h(x_1, \dots, x_n)$$

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$

$$f(x_1, \ldots, x_n | p) = p^{\sum_i x_i} (1-p)^{n - \sum_i x_i}$$

$$= \underbrace{g\left(\textstyle\sum_i x_i \,|\, p\right)}_{p^{\sum_i x_i} (1-p)^{n-\sum_i x_i}} \underbrace{h(x_1, \ldots, x_n)}_{=1}$$