

STATS300B – Lecture 8

Julia Palacios
Scribed by Michael Howes

01/27/22

Contents

1 Fisher information	1
-----------------------------	----------

1 Fisher information

Suppose we have a model $\mathcal{P} = \{P_\theta : \theta \in \Omega\}$ and $\Omega \subseteq \mathbb{R}^d$. Define $l_\theta(x) = \log(p_\theta(x))$ where p_θ is the density of P_θ . The *Fisher information of θ* is defined to be the matrix,

$$I(\theta) = \mathbb{E}[\nabla l_\theta(X) \nabla l_\theta(X)^T] = \text{Var}_\theta(\nabla l_\theta(X)).$$

If integration and differentiation can be exchanged, then

$$I(\theta) = -\mathbb{E}[\nabla^2 l_\theta(X)].$$

Under appropriate assumptions about the regularity of the map $\theta \mapsto l_\theta$, then for all θ in the interior of Ω ,

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}_k(0, I(\theta)^{-1}),$$

where $\hat{\theta}_n$ is the MLE from a sample $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} P_\theta$.