Estimation, Exponential families and Equivariance

Gunnar.Taraldsen@ntnu.no
Department of Mathematical Sciences
Norwegian University of Science and Technology

September 18, 2022

Abstract

Exponential families and equivariance are central concepts in theoretical statistics. This is demonstrated by solving the problems formulated in this note. Some knowledge of probability theory, statistics, and multivariate calculus is assumed. You must, as always, formulate necessary assumptions as part of the proof of a claim. Casella and Berger (2002) give many more illustrations and explanations of the theory behind the methods of statistical inference. [1]

Contents

---

[1] 2020 Mathematics Subject Classification: 62-01 Introductory exposition pertaining to statistics; 62F10 Point estimation; 62B05 Sufficient statistics and fields; 62C05 General considerations in statistical decision theory; 58K70 Symmetries, equivariance on manifolds; 62E10 Characterization and structure theory of statistical distributions;

1 Sufficient Estimators in Exponential Families

1.1 Exponential Families

Assume that

$$f(x \mid \theta) = \exp(\theta t - \gamma) h(x) \tag{1}$$

is the density of a random quantity X given by a real statistic $t = t(x)$ and a real model parameter $\theta \in \Omega_\Theta \subset \mathbb{R}$. Determine $\gamma = \gamma(\theta)$ and find the largest model space Ω_Θ which is consistent with equation (1). Can you say something about convexity or other properties of Ω_Θ? Show that $T = t(X)$ is the UMVUE and MLE for $\tau = \mathrm{E}(T)$.

1.2 Important Moments

Use the moment generating function of T in equation (1) to prove that $\tau = \gamma'$, and that τ and θ are in a one-one correspondence. Can this be used to prove existence and uniqueness of the MLE $\hat{\theta}$? What does the Cramer-Rao inequality imply for $\hat{\tau}$ and $\hat{\theta}$? Show that the density of t is on the form given by the special case $t = x$. Use this to give a proof of completeness of t.

1.3 The Classics

Consider X to be a random sample from $\mathbf{N}(\mu, \sigma^2)$ with σ known. Is this a special case of equation (1)? If so give the corresponding results from the analysis in section 1.1. Do the same for the cases where μ is known. Consider the same analysis for all six cases given by $\mathbf{Gamma}(\alpha, \beta)$, $\mathbf{B}(n, p)$, and $\mathbf{U}((1-k)\beta, (1+k)\beta)$ where $0 < k \leq 1$.

1.4 Location and Scale

Prove that $Y > 0$ is from a scale family if and only if $X = \ln Y$ is from a location family. Prove that $Y > 0$ is from a exponential family if and only if X is. Determine for each of the cases in section 1.3 if X or T are from a location family or a scale family or not. What about invariance for other transformations for the examples in section 1.3? Can you give conditions on equation (1) that ensures that X or T is from a location family or a scale family?

1.5 Exponential Families in Space

Redo the exercise in Section 1.1 for the case where $t, \theta \in \mathbb{R}^k$. Analyze in particular the case where X is random sample from $\mathbf{N}(\mu, \sigma^2)$ with both parameters unknown. Do the same analysis for the $\mathbf{Gamma}(\alpha, \beta)$ case. What happens if there is a restriction $M(\theta) = m$ on the model parameter θ? Consider in particular the example given by $M(\mu, \sigma) = \mu - \sigma = 0$ for the $\mathbf{N}(\mu, \sigma^2)$ case.

2 Optimal Equivariant Estimators

2.1 The Basu Theorem

Prove that

$$\mathrm{E}(\mathrm{E}(\phi(S) \,|\, T) - \mathrm{E}(\phi(S))) = 0 \tag{2}$$

and use this to prove the Basu theorem.

2.2 Scale Equivariance

Let the data x be a positive number. Prove that $t = t(x) = xt_1$ is a scale equivariant estimator, and that all scale equivariant estimators are on this form for $x, t_1 > 0$. What about location equivariant estimators for real data?

2.3 Invariant Loss

Let a loss for an estimator t be defined by $\ell = \ell(t, \beta) = \phi(t\beta^{-1})$ with $\phi(y) = [\ln(y)]^2$. Show that ℓ is invariant in the sense that $\ell(gt, g\beta) = \ell(t, \beta)$. Can you determine all possible scale invariant loss functions? What about location invariant loss functions for real data? Convex and invariant loss?

2.4 Group Model

Show that the statistical model for the data $X = \beta U$ is scale invariant where $U > 0$ has a known distribution. Explain that the identity $\ell(t(\beta U), \beta) = \ell(t(x), xU^{-1})$ gives that the risk $\rho = \mathrm{E}(\ell(T, \beta))$ is minimized by the statistic $t = \exp(\mathrm{E}(\ln(xU^{-1})))$. Calculate t for the case where x is the minimal sufficient statistic from a random sample from the $\mathbf{Gamma}(\alpha, \beta)$ with known shape α. Do this also for the case $\mathbf{N}(\mu, \sigma^2)$ where μ is known. What if the loss is defined using $\phi(y) = y^2 - 1 - \ln(y^2)$ as Stein did?

2.5 Conditionality Principle?

Show that $w = \overline{x}$ is scale equivariant. Is the statistic $a = (x_1, x_2)w^{-1}$ ancillary for data $X = (X_1, X_2) = (\beta U_1, \beta U_2)$? Is the optimal solution for the conditional model for w given $A = a$ also unconditionally optimal? Use the Basu theorem to prove consistency between the solutions by this method compared to the method of sufficient statistics used in section 2.4. Draw the level sets of the statistic a and the sufficient statistic and explain the kind of data reduction obtained by the two methods.

References

Casella, G. and R. L. Berger (2002). *Statistical Inference* (2nd edition ed.). Duxbury, Thomson learning.