Source: http://www2.uah.es/estaeco/Estadistica/Dtee/Neyman_Fisher_Theorem.PDF

Neyman-Fisher, Theorem

Better known as "Neyman-Fisher Factorization Criterion", it provides a relatively simple procedure either to obtain sufficient statistics or check if a specific statistic could be sufficient.

Fisher was the first who established the Factorization Criterion like a sufficient condition for *sufficient* statistics in 1922. Years later, Neyman demonstrated it necessity under certain restrictive conditions in 1935. Finally, Halmos and Savage extended it in 1949 as follows:

> Let $\wp=\{P_\theta,\theta\in\Omega\}$ be a family of probability measures on a measurable space $(\Theta_X,\mathcal{A})$ absolutely continuous with respect to a σ-finite measure μ. Let suppose that its probability densities in the Radon-Nicodym sense $p_\theta = dP_\theta/d\mu$ exist a.s.[μ] (almost sure for μ).
> A necessary and sufficient condition for the sufficiency with respect to $\wp$ of a statistic T transforming the probability space $(\Theta_X,\mathcal{A},P_\theta)$ into $(\Theta_T,\mathcal{B},P_T)$ is the existence $\forall\theta\in\Omega$ of a $T^{1}(\mathcal{B})$-measurable function $g_\theta T(x)$ and an $\mathcal{A}$-measurable function $h(x)\neq 0$ a.s.[P_θ], both defined $\forall x\in\Theta_X$, non-negatives and μ-integrable, such that $p_\theta(x)= g_\theta T(x)\cdot h(x)$, a.s.[$\mu$].

Densities p_θ can be either probability density functions from absolutely continuous random variables or probability functions from discrete random variables among other possibilities, depending on the nature and definition of μ.

In common economic practice, this Factorization Criterion adopts simpler appearances. Thereby, in Estimation Theory under random sampling, the criterion is usually enounced as follows:

> Let X be a random variable belonging to a regular family of distributions $F(x;\theta)$ which depends on a parameter θ (mixture of absolutely continuous and discrete random variables on values not depending on the parameter) representing some characteristic of certain population. Moreover, let $x=(X_1, X_2, \ldots, X_n)$ represent a random sample size n of X, extracted from such a population.
> A necessary and sufficient condition for the sufficiency of a statistic $T=t(x)$ with respect to the family of distributions $F(x;\theta)$ is that the sample likelihood function $L_n(x;\theta)$ could be factorized like $L_n(x;\theta)=g(T;\theta)\cdot h(x)$. Here, "g" and "h" are non-negative real functions, "g" depending on sample observations through the statistic exclusively and "h" not depending on the parameter.

When the random variable is absolutely continuous (discrete), function $g(t;\theta)$ is closely related to the probability density function (probability function) of the statistic T. Thus, the criterion could be equivalently enounced assuming the function $g(t;\theta)$ to be exactly such probability density function (probability function). In this case, the usually complex work of deducing the distribution $x|T$ of the original observations conditioned to the statistic T becomes easier, being specified by h(x).

According to this Factorization Criterion, any invertible function of a sufficient statistic $T^*=k(T)$ is a sufficient statistic too.

Likewise, exponential family defined by probability density functions like $f(x;\theta)=k(x)\cdot p(\theta)\cdot \exp[c(\theta)'T(x)]$ always admits a sufficient statistic T.

<u>References</u>:

Fisher,R.A. (1922) "On the mathematical foundation of theoretical statistics" *Philos. Trans. Roy. Soc. serie A.*, 222:309-368.
Neyman,J. (1935) "Sur un teorema concernente le cosidette statistiche sufficienti" *Giorn. Ist. Ital. Att.,* 6:320-334.
Halmos,P.R. and Savage,L.J. (1949) "Application of the Radon-Nikodym theorem to the theory of sufficient statistics" *Ann. Math. Statist.*, 20:225-241.
Zacks,S. (1971) *The Theory of Statistical Inference.* John Wiley & Sons, Inc. 38-48.