Lecture 42: Empirical distribution function; probability integral transformation

Statistical Inference–II ▪ Units IV–V: Nonparametric Inference

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

M.Sc. Statistics, Semester III
Academic Session 2026–2027

This lecture

Previous lecture

Lecture 41: Nonparametric and distribution-free procedures; assumptions that still matter.

Today

- Empirical Distribution Function and Probability Integral Transformation
  - Probability Integral Transformation
  - Importance in Nonparametric Inference

The empirical distribution function

Let $X_1, \ldots, X_n$ be a random sample from a distribution function F. The *empirical distribution function* (e.d.f.) is

$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(X_i \leq x).$$

It is a right-continuous step function that places probability $1/n$ at each observation. For every fixed x,

$$\mathbb{E}\{\widehat{F}_n(x)\} = F(x), \qquad \mathrm{Var}\{\widehat{F}_n(x)\} = \frac{F(x)\{1 - F(x)\}}{n}.$$

The e.d.f. is the basic estimator used in Kolmogorov–Smirnov procedures.

The quantile function

For a distribution function F and $0 < u < 1$ define the *quantile function*

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}.$$

When F is continuous and strictly increasing this is the ordinary inverse function; the definition with the infimum also covers flat stretches and jumps of F.

Probability integral transformation

Theorem 42.1 (Probability integral transformation)

If X has a continuous distribution function F, then $U = F(X)$ has the uniform distribution on $(0, 1)$.

Proof.

We give the proof when F is continuous and strictly increasing on the set where $0 < F < 1$; then F^{-1} is the ordinary inverse and $F\{F^{-1}(u)\} = u$. Because F is increasing, $F(X) \leq u$ exactly when $X \leq F^{-1}(u)$. Hence, for $0 < u < 1$,

$$\mathbb{P}(U \leq u) = \mathbb{P}\{F(X) \leq u\} = \mathbb{P}\{X \leq F^{-1}(u)\} = F\{F^{-1}(u)\} = u.$$

Since $0 \leq U \leq 1$, $U \sim U(0, 1)$. (Continuous F with flat stretches: full proof in Lecture Notes, Appendix A.) □

Discontinuous F: the randomised transformation

Remark 42.2

Continuity is essential for the simple statement. If F has jumps, $F(X)$ is not generally uniform. A randomised version,

$$U = F(X^-) + V\{F(X) - F(X^-)\},$$

where $V \sim U(0,1)$ is independent of X, is uniform even for a discontinuous F.

Inverse probability transformation

Theorem 42.3 (Inverse probability transformation)

If $U \sim U(0,1)$, then $X = F^{-1}(U)$ has distribution function F. This holds for every distribution function F, continuous or not.

Proof.

For a continuous, strictly increasing F, $F^{-1}(U) \le x$ exactly when $U \le F(x)$, so

$$\mathbb{P}(X \le x) = \mathbb{P}\{F^{-1}(U) \le x\} = \mathbb{P}\{U \le F(x)\} = F(x),$$

since $\mathbb{P}(U \le c) = c$ for $0 \le c \le 1$. The key equivalence $F^{-1}(u) \le x \iff u \le F(x)$ holds for *every* distribution function (Lecture Notes, Appendix A), so the same one-line argument proves the general case. □

Example: the exponential distribution

Example 42.4 (Exponential distribution)

Let X have exponential distribution with rate λ, $F(x) = 1 - e^{-\lambda x}$ for $x > 0$. Then

$$U = 1 - e^{-\lambda X} \sim U(0, 1).$$

Conversely, if $U \sim U(0, 1)$, then

$$X = -\frac{1}{\lambda} \log(1 - U)$$

has the exponential distribution. Since $1 - U$ is also uniform, one may use $X = -\lambda^{-1} \log U$ for simulation.

Importance in nonparametric inference

Under a fully specified continuous null distribution F_0, the transformed observations

$$U_i = F_0(X_i)$$

are i.i.d. uniform.

Therefore any statistic built from the relative discrepancy between the e.d.f. of the U_i's and the uniform distribution has a null distribution independent of F_0.

This is the conceptual basis of the one-sample Kolmogorov–Smirnov test.

After this lecture

Practice (Lecture Notes, Chapter 3)

- Short-answer questions (Section 3.14): Q2.
- Exercises (Section 3.15): 3.2, 3.3.

Next lecture

Lecture 43: Order statistics; the distribution of $X_{(r)}$.