STSCI 5080
Probability Models and Inference
Lecture 10: Order Statistics and Expected Values

September 27, 2018

Independence

Definition

Random variables $X_1, \ldots, X_n$ are independent if

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \cdots P(X_n \in A_n)$$

for any subsets $A_1, \ldots, A_n \subset \mathbb{R}$.

Definition

If random variables $X_1, \ldots, X_n$ are independent with common cdf F (on $\mathbb{R}$), then they are called a random sample from F.

$$\text{“}X_1, \ldots, X_n \sim F \text{ i.i.d.”},$$

where “i.i.d.” means “independent and identically distributed”. E.g.,

$$X_1, \ldots, X_n \sim N(\mu, \sigma^2) \text{ i.i.d.}$$

Distributions of maximum and minimum

Example

If $X_1, \ldots, X_n$ are i.i.d. with common cdf F and pdf f, then find the cdfs and pdfs of $U = \max_{1 \leq i \leq n} X_i$ and $V = \min_{1 \leq i \leq n} X_i$.

Theorem

The cdf of $U = \max_{1 \le i \le n} X_i$ *is*

$$F_U(u) = \{F(u)\}^n.$$

Hence, the pdf of U *is*

$$f_U(u) = nf(u)\{F(u)\}^{n-1}.$$

Proof.

Since

$$\max_{1 \le i \le n} X_i \le u \Leftrightarrow X_i \le u \text{ for all } 1 \le i \le n,$$

we have

$$\begin{aligned} F_U(u) &= P\left(\max_{1 \le i \le n} X_i \le u\right) = P(X_1 \le u, \ldots, X_n \le u) \\ &= P(X_1 \le u) \cdots P(X_n \le u) = \{F(u)\}^n. \end{aligned}$$

$\square$

Example 10.1

Example

Let X and Y be lifetimes (in year) of two cars, and suppose that $X, Y \sim Ex(\lambda)$ i.i.d. with $\lambda = 0.1$. What is the probability that at least one car will be working for more than 10 years?

Example 10.1

Example

Let X and Y be lifetimes (in year) of two cars, and suppose that $X, Y \sim Ex(\lambda)$ i.i.d. with $\lambda = 0.1$. What is the probability that at least one car will be working for more than 10 years?

The cdf of $Ex(0.1)$ is $F(x) = 1 - e^{-0.1 \cdot x}$ for $x \geq 0$. The cdf of $U = \max\{X, Y\}$ is

$$F_U(u) = (1 - e^{-0.1 \cdot u})^2$$

for $u \geq 0$, so that

$$P(U > 10) = 1 - F_U(10) = 1 - (1 - e^{-1})^2 \approx 0.60.$$

Theorem

The cdf of $V = \min_{1 \le i \le n} X_i$ *is*

$$F_V(v) = 1 - \{1 - F(v)\}^n.$$

Hence, the pdf of V *is*

$$f_V(v) = nf(v)\{1 - F(v)\}^{n-1}.$$

Proof.

Since

$$\min_{1\le i\le n} X_i > v \Leftrightarrow X_i > v \text{ for all } i = 1,\ldots,n,$$

we have

$$\begin{aligned} P\left(\min_{1\le i\le n} X_i > v\right) &= P(X_1 > v,\ldots,X_n > v) \\ &= P(X_1 > v)\cdots P(X_n > v) = \{1 - F(v)\}^n. \end{aligned}$$

Hence, $F_V(v) = 1 - P\left(\min_{1\le i\le n} X_i > v\right) = 1 - \{1 - F(v)\}^n$. $\square$

Example 10.2

Example

If $X_1, \ldots, X_n \sim Ex(\lambda)$ i.i.d., find the distribution of $V = \min_{1 \le i \le n} X_i$.

Example 10.2

Example

If $X_1, \ldots, X_n \sim Ex(\lambda)$ i.i.d., find the distribution of $V = \min_{1 \le i \le n} X_i$.

Since $F(x) = 1 - e^{-\lambda x}$ for $x \ge 0$, we have

$$F_V(u) = 1 - \{1 - F(v)\}^n = 1 - e^{-n\lambda v}$$

for $v \ge 0$. This is the cdf of $Ex(n\lambda)$, and so $V \sim Ex(n\lambda)$.

Statistic

Definition

Let $X_1, \ldots, X_n$ be a random sample. Then a function of $X_1, \ldots, X_n$ is called a statistic of $X_1, \ldots, X_n$

Order statistics

Definition

Let $X_1, \dots, X_n$ be a random sample from a pdf f. Then the order statistics $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ are defined by sorting $X_1, \dots, X_n$ in increasing order.

Example

$X_{(1)} = \min_{1 \le i \le n} X_i$ and $X_{(n)} = \max_{1 \le i \le n} X_i$.

Densities of order statistics

Theorem

Let $X_1, \ldots, X_n$ be independent continuous random variables with common pdf f, and let $X_{(1)} < \cdots < X_{(n)}$ be order statistics of $X_1, \ldots, X_n$. For any $k = 1, \ldots, n$, the pdf of $X_{(k)}$ is

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} f(x) F(x)^{k-1} \{1 - F(x)\}^{n-k},$$

where F denotes the cdf corresponding to f.

See the supplemental material for a proof.

Chapter 4 Expected Values

Expected values

- A random variable may take many different values.
- The expected value/expectation/mean is a summary of the possible values of a random variable. “average value”.

Expectation: discrete case

Definition

The expected value (or expectation or mean) of a discrete random variable X with pmf $p(x)$ is defined by

$$E(X) = \sum_x xp(x)$$

whenever the sum is absolutely convergent, i.e., $\sum_x |x|p(x) < \infty$.

If X is non-negative, then we also define

$$E(X) = \sum_x xp(x) = \sum_{x \geq 0} xp(x)$$

even when $\sum_{x \geq 0} xp(x) = \infty$.

Example 10.3

Example

Find the expected values of the following random variables: (a) X is a Bernoulli random variable with success probability p. (b) $X \sim P(\lambda)$.

Example 10.3

Example

Find the expected values of the following random variables: (a) X is a Bernoulli random variable with success probability p. (b) $X \sim P(\lambda)$.

(a) $E(X) = p$. (b) $E(X) = \lambda$.

Some cautions

- The expectation is a non-random number.
- The expectation is determined uniquely by the pmf or the cdf.
- “The Poisson distribution $Po(\lambda)$ has mean λ.”

Example 10.4

Example

If X takes values in $2^k, k = 0, 1, , \ldots$ such that

$$P(X = 2^k) = \frac{1}{2^{k+1}},$$

then

$$E(X) = \sum_x xP(X = x) = \sum_{k=0}^{\infty} 2^k \frac{1}{2^{k+1}} = \infty.$$

Expectation: continuous case

Definition

The expected value (or expectation or mean) of a continuous random variable X with pdf $f(x)$ is defined by

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

whenever $\int_{-\infty}^{\infty} |x|f(x)dx < \infty$.

If X is non-negative, then we also define

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{\infty} xf(x)dx$$

even when $\int_{0}^{\infty} xf(x)dx = \infty$.

Example 10.5

Example

Find the expected values of the following random variables: (a) $X \sim Ex(\lambda)$. (b) $X \sim N(\mu, \sigma^2)$.

Example 10.5

Example

Find the expected values of the following random variables: (a) $X \sim Ex(\lambda)$. (b) $X \sim N(\mu, \sigma^2)$.

(a) $E(X) = \frac{1}{\lambda}$. (b) $E(X) = \mu$.

Example 10.6

Example

If X has the Cauchy density

$$f(x) = \frac{1}{\pi(1+x^2)}, \; -\infty < x < \infty,$$

then $E(X)$ is not well-defined.

Example 10.6

Example

If X has the Cauchy density

$$f(x) = \frac{1}{\pi(1+x^2)}, \ -\infty < x < \infty,$$

then $E(X)$ is not well-defined.

X takes both positive and negative values, but

$$\int_{-\infty}^{\infty} |x| \frac{1}{\pi(1+x^2)} dx = \frac{2}{\pi} \int_0^{\infty} \frac{x}{1+x^2} dx = \frac{1}{\pi} \left[\log(1+x^2)\right]_0^{\infty} = \infty.$$