Math 30530, Introduction to Probability

Spring 2019

Notes on continuous random variables

Continuous random variables: A random variable X is *continuous* if

1 it takes values in a continuum (all the real line, or some subinterval of the real line, or a union of subintervals of the real line)

(so, the set of possible outcomes of the experiment that leads to X is not discrete, or listable), and also if

2 there is a function $f : \mathbb{R} \to [0, \infty)$ such that for each $x \in \mathbb{R}$ we can model (or calculate) the probability that X takes a value less than or equal to X via an integral of f, specifically:

$$P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt.$$

Some facts:

- f is called the *density function*, or *probability density function*, or *pdf*, of X.
- Since $P(X \in \mathbb{R}) = 1$ we have $\int_{-\infty}^{\infty} f(t)\, dt = 1$.
- The function $F_X : \mathbb{R} \to [0, 1]$ defined by $F_X(x) = P(X \leq x)$ is called the *distribution function*, or *cumulative distribution function*, or *cdf*, of X. By a theorem of calculus, it is a continuous function, that is 0 at $-\infty$, is increasing on $\mathbb{R}$, and is 1 at ∞.
- The distribution function can be obtained from the density function by integration; by the fundamental theorem of calculus, the density function can be found from the distribution function by differentiation:

$$f(x) = \frac{d}{dx} F_X(x).$$

- For any continuous random variable, $P(X = c) = 0$ for any particular number c, and so there is no difference between

$$\begin{array}{ll} P(a < X < b) & P(a < X \leq b) \\ P(a \leq X < b) & P(a \leq X \leq b). \end{array}$$

  Each of these can be computed as $\int_a^b f(t)\, dt$.

- The expectation of X is calculated as

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

and the variance is calculated as $\mathrm{Var}(X) = E((X - E(X))^2)$.

- The law of the unconscious statistician for continuous random variables says that for any function g,

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\ dx.$$

All the consequences of the law of the unconscious statistician discrete random variables hold also for continuous random variables, in particular

Monotonicity if $f(x) \leq g(x)$ for all X then $E(f(x)) \leq E(g(X))$.

Linearity $E(af(X) + bg(X)) = aE(f(X)) + bE(g(X))$.

Variance formula $\mathrm{Var}(X) = E(X^2) - E(X)^2$.

Markov's inequality If $X \geq 0$ always, for any $t > 0$,

$$P(X \geq tE(X)) \leq \frac{1}{t}.$$

Chebychev's inequality For any X and any $c > 0$,

$$P(|X - E(X)| \geq c\sigma) \leq \frac{1}{c^2}.$$

The five basic continuous random variables: Here is a list of the five commonly occurring continuous random variables. For each, I've listed its name, abbreviation, typical usage, parameters, range of possible values, density function, expectation and variance. I've also added a few special properties of some of the random variables.

Uniform Uni($[a, b]$) or Uniform($[a, b]$) — Models the selection of a real number uniformly at random from the interval $[a, b]$

- **Parameters**: real numbers $a < b$
- **Range**: $[a, b]$
- **Density function**:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

- **Expectation**: $\frac{a+b}{2}$
- **Variance**: $\frac{(b-a)^2}{12}$

Exponential Exp(λ) or Exponential(λ) — Models waiting time until the first occurrences of a rare event, or time between consecutive occurrences, when there are on average λ occurrences per unit time (the *rate* of occurrences is λ), when occurrences in disjoint time periods are independent, and when simultaneous occurrences are very unlikely (so — same λ as in Poisson random variable)

- **Parameters**: $\lambda > 0$
- **Range**: $(0, \infty)$
- **Density function**:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

- **Expectation**: $\frac{1}{\lambda}$
- **Variance**: $\frac{1}{\lambda^2}$
- **Special property**: The exponential is the only continuous random variable that is *memoryless*: for $s, t > 0$,

$$P(X > s + t | X > t) = P(X > s)$$

Standard Normal Z or $N(0, 1)$ — The basic building block distribution of the general normal (see below)

- **Parameters**: No varying parameters
- **Range**: $\mathbb{R}$
- **Density function**:

$$f(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

- **Expectation**: 0
- **Variance**: 1
- **Special property**: Always denoted Z, and cumulative distribution function always denoted Φ, so

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt$$

General Normal $N(\mu, \sigma^2)$ — Models the result of aggregating a lot of independent random factors, when the average value of the aggregate is μ and the variance is σ^2

- **Parameters**: $\mu \in \mathbb{R}$, $\sigma^2 > 0$
- **Range**: $\mathbb{R}$
- **Density function**:

$$f(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

- **Expectation**: μ
- **Variance**: σ^2
- **Special property**: $N(\mu, \sigma^2) = \sigma Z + \mu$, which means that the *z-score* of $N(\mu, \sigma^2)$, $(x - \mu)/\sigma$, is a standard normal

Gamma $\Gamma(n, \lambda)$ — Different values of n and λ give different shapes of density function, making it a versatile random variable for modeling; also, some special values of n and λ give some meaningful distributions (see below)

- **Parameters**: $n > 0$ (not necessarily an integer), $\lambda > 0$; n is shape parameter, as changing it fundamentally changes the shape of the density function, and λ is the rate parameter, explained below
- **Range**: $(0, \infty)$
- **Density function**:

$$f(x) = \begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

  where $\Gamma(n) = \int_0^\infty t^{x-1} e^{-t}\, dt$
- **Expectation**: $\frac{n}{\lambda}$
- **Variance**: $\frac{n}{\lambda^2}$
- **Special properties**:
  - If $n = 1$, $\Gamma(1, \lambda) \sim \text{Exp}(\lambda)$.
  - If n a whole number, $\Gamma(n, \lambda)$ models the waiting time to the nth occurrence of an event, when the rate of occurrences is λ. In this sense

    Exponential is to Geometric as Gamma is to Negative Binomial
  - If Z is a standard normal, then $Z^2 \sim \Gamma(1/2, 1/2)$.
  - If $Z_1, Z_2, \ldots, Z_n$ are n independent readings from a standard normal, then $Z_1^2 + \cdots + Z_n^2 \sim \Gamma(n/2, 1/2)$ — this is called the *chi-squared distribution with n degrees of freedom*