Probability and Statistics Notes

Chapter Three

Jesse Crawford

Department of Mathematics
Tarleton State University

Outline

1. Section 3.1: Continuous Data
2. Section 3.2: Continuous Random Variables

Discrete Data

Possible Values

- Countable.
- Typically either
  - $\{0, 1, \ldots, n\}$ or
  - $\{0, 1, 2, \ldots\}$

Example

- Number of free throws made in a sequence of 10 attempts.
- Possible values: $0, 1, \ldots, 10$
- Typical sample: $9, 10, 7, 10, 9, 6, 7, 8$

Example

- Number of phone calls received in a five minute period.
- Possible values: $0, 1, 2, \ldots$
- Typical sample: $28, 28, 16, 29, 28, 20, 24, 15$

Continuous Data

Possible Values

- Not countable
- Usually some interval of real numbers

Example

- Weights of randomly selected men.
- Set of possible values: $(0, \infty)$
- or perhaps $(100, 190)$
- Typical sample:
  $161.34, 151.06, 137.38, 136.99, 131.95, 134.49, 140.86, 110.86$

Example

- Measurement errors
- Set of possible values: $(-\infty, \infty)$
- or perhaps $(-1, 1)$
- Typical sample: $-0.407, -0.398, 0.059, 0.555$

Outline

1. Section 3.1: Continuous Data

2. Section 3.2: Continuous Random Variables

Definition

X is a *continuous random variable* if there is a function $f : \mathbb{R} \to \mathbb{R}$ such that

- $f(x) \geq 0$, for all $x \in \mathbb{R}$,
- $\int_{-\infty}^{\infty} f(x)dx = 1$, and
- $P(a \leq X \leq b) = \int_a^b f(x)dx$.

f is the *probability density function* (p.d.f.) of X.

Calculating with Continuous Random Variables

- $E(X) = \int_{-\infty}^{\infty} x f(x) dx$
- $E(u(X)) = \int_{-\infty}^{\infty} u(x) f(x) dx$
- $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$
- $\text{Var}(X) = E(X^2) - E(X)^2$
- $\sigma_X = \sqrt{\text{Var}(X)}$
- $M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$

Cumulative Distribution Function (c.d.f.)

- Cumulative distribution function: $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$.
- $P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x)dx$.
- $F'(x) = f(x)$, if f is continuous at x.
- Properties of F:
  - F is nondecreasing
  - $\lim_{x\to\infty} F(x) = 1$
  - $\lim_{x\to-\infty} F(x) = 0$
  - If X is a continuous random variable, F is continuous.
- Note that f need not be continuous, and in fact f can be modified at finitely many points, and it will still be a p.d.f. for X.

Review for Exam 3

- Section 3.1: Continuous Data
  - Relative frequency histogram for continuous data
  - Sample percentiles/quantiles/quartiles
- Section 3.2: Continuous Random Variables
  - p.d.f.
  - $\sum \to \int$
  - c.d.f.
  - Theoretical percentiles/quantiles/quartiles
- Section 3.3: Uniform and Exponential Distributions
  - Uniform distribution

$$f(x) = \frac{1}{b-a}, \text{ for } a \leq x \leq b$$

$$F(x) = \frac{x-a}{b-a}, \text{ for } a \leq x \leq b$$

  - $\mu = \frac{a+b}{2}$

- Section 3.3: Uniform and Exponential Distributions
  - Exponential Distribution

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \text{ for } x > 0$$

$$F(x) = 1 - e^{-\frac{x}{\theta}}, \text{ for } x > 0$$

  - $\mu = \theta$ and $\sigma^2 = \theta^2$
  - Waiting time until 1st phone call in a Poisson process.
- Section 3.4: The Gamma and Chi-square Distributions
  - Gamma distribution with parameters $\alpha, \theta > 0$

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\theta}}, \text{ for } x > 0$$

  - $\mu = \alpha\theta$ and $\sigma^2 = \alpha\theta^2$
  - Waiting time until αth phone call in a Poisson process.
  - $\chi^2(r) = \text{Gamma}(\alpha = \frac{r}{2}, \theta = 2)$
  - Be able to find $\chi^2_{0.05}(r)$ etc.

- Section 3.5: Distributions of Functions of a Random Variable
  - Given X, find p.d.f. of $Y = u(X)$
  - Distribution function technique

$$G(y) = P(Y \leq y) = P[u(X) \leq y]$$

$$g(y) = G'(y)$$

  - Change of variables technique. Only works if u is increasing or decreasing.

$$v = u^{-1}$$

$$g(y) = f[v(y)]|v'(y)|$$