4 Continuous Random Variables and Their Probability Distributions

4.1 Introduction
- Not all random variables of interest are discrete. E.g. daily rainfall; life expectancy in years;
- A random variable that takes on any value in an interval is called “continuous”: no values between a and b can be ruled out as possible value
- The probability distribution for a continuous RV cannot be specified easily as discrete RVs. Cannot assign nonzero probabilities to all the points on a line interval and making them sum to 1

4.2 The probability distribution for a continuous random variable
- Definition 4.1 The “(cumulative) distribution function” \( F(y) \) of a random variable \( Y \) is given by \( F(y) = P(Y \leq y) \) for \(-\infty < y < \infty\).
- Example 4.1 What’s \( F(y) \) of \( Y \sim \text{binom}(2, 1/2) \)? A graph? A “step function”
- Distribution functions for discrete RVs are always step functions since the function increases only at a countable # of points
- **Theorem 4.1** If \( F(y) \) is a distribution function, then
  1. \( F(-\infty) = 0 \)
  2. \( F(\infty) = 1 \)
  3. \( F(y) \) is non-decreasing function of \( y \).
- Definition 4.2 A RV \( Y \) is said to be “continuous” if its distribution function \( F(y) \) is continuous for \(-\infty < y < \infty\).
  - For a continuous RV \( Y \), we must have \( P(Y = y) = 0 \). (Why?)
- **Definition 4.3** Let \( F(y) \) be the distribution function for a continuous RV \( Y \). Then \( f(y) \), given by \( f(y) = \frac{dF(y)}{dy} = F'(y) \) whenever the derivative exists, is called the “probability density function (PDF)” for the RV \( Y \).
- It follows that \( F(y) = \int_{-\infty}^{y} f(t)dt \) (graph?)
- **Theorem 4.2** If \( f(y) \) is a density function, then
  1. \( f(y) \geq 0 \) for any values of \( y \).
  2. \( \int_{-\infty}^{\infty} f(y) = 1 \).
- Example 4.2 \( F(y) = y \) \( 1(0 \leq y < 1) + 1(y \geq 1) \). What’s the PDF for \( Y \)? Graph it.
Example 4.3 Let Y be a continuous RV with PDF $f(y)=3y^2$ if $0<y<1$. Find $F(y)$. Graph both $f(y)$ and $F(y)$.

$F(y_0)=P(Y\leq y_0)$. What about $P(a\leq Y\leq b)$? Recall the frequency distribution situation: “area under the frequency distribution”.

**Theorem 4.3 For $a\leq b$,**

\[ P(a \leq Y \leq b) = P(a < Y < b) = P(a < Y < b) = \int_{a}^{b} f(y)dy . \]

Example 4.4, 4.5. Given $f(y)=cy^2$ if $0<y<2$, find the value of $c$ for which $f(y)$ is a valid density function. Also find $P(1<Y<2)$.

Some density functions provide good models for population frequency distributions encountered in nature.

**Exercises 4.1~13**

**Keywords: continuous RV, PDF**

### 4.3 Expected values for continuous random variables

- **Definition 4.4 (The expected value of a continuous RV)**
  \[ E(Y) = \int_{-\infty}^{\infty} yf(y)dy \]

- **Theorem 4.4 (The expected value of a function of Y)**
  \[ E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy \]

- **Theorem 4.5** For any constant $c$, and functions $g(Y)$, $g_1(Y)$, $g_2(Y)$, ..., $g_k(Y)$ of continuous $Y$,
  
  1. $E(c)=c$.
  2. $E[cg(Y)] = cE[g(Y)]$.
  3. $E[g_1(Y)+g_2(Y)+...+g_k(Y)]=E[g_1(Y)]+E[g_2(Y)]+...+E[g_k(Y)]$.

Example 4.6 For example 4.4, find $\mu=E(Y)$ and $\sigma^2=V(Y)$.

**Exercises 4.14~27**

**Keywords: E(Y), E(g(Y))**

### 4.4 The uniform probability distribution

- Probability (the bus will arrive in any given subinterval of time) $\propto$ the length of the subinterval:
  an example of

- **Definition 4.5 If $a<b$, a RV Y is said to have a continuous “uniform probability distribution” on the interval $(a, b)$ iff**
  \[ f(y) = \frac{1}{b-a}1(a < y < b) . \]

- **Definition 4.6 The constant that determine the specific form of a density function is called “parameters” of the density function**
Example 4.7 Arrivals of customers at a checkout counter ~ Poisson. It is known that during a given 30-minute period, one customer arrived at the counter. \( P(\text{the customer arrived during the last 5 min of the 30-minute period}) = ? \) [Hint: It follows Uniform(0,30)]

Uniform distribution is critical for simulation study: can generate \( y \sim F(y) \) by transforming a uniform RV.

**Theorem 4.6** If \( Y \sim \text{Uniform}(a, b) \), then
\[
\mu = E(Y) = \frac{b - a}{2} \quad \text{and} \quad \sigma^2 = V(Y) = \frac{(b - a)^2}{12}
\]

**Exercises 4.28~45**

**Keywords**: Uniform(\( a, b \))

### 4.5 The normal probability distribution

- The most widely used continuous distribution; the “bell” curve;
- There’re reasons why we see them so often: the “central limit theorem” (Chapter 7)
- **Definition 4.7** A RV \( Y \) is said to have a “normal probability distribution” or \( Y \sim \text{N}(\mu, \sigma^2) \) iff
\[
f(y) = \frac{1}{\sigma \sqrt{2 \pi}} \exp \left( -\frac{(y - \mu)^2}{2\sigma^2} \right) \quad \text{for} \ \sigma > 0 \ \text{and} \ \mu.
\]

Example 4.7 Arrivals of customers at a checkout counter ~ Poisson. It is known that during a given 30-minute period, one customer arrived at the counter. \( P(\text{the customer arrived during the last 5 min of the 30-minute period}) = ? \) [Hint: It follows Uniform(0,30)]

Uniform distribution is critical for simulation study: can generate \( y \sim F(y) \) by transforming a uniform RV.

**Theorem 4.6** If \( Y \sim \text{N}(\mu, \sigma^2) \), then
\[
\mu = E(Y) \quad \text{and} \quad \sigma^2 = V(Y)
\]

- **Proof**: use MGF (section 4.9)

- Evaluating the integral can only be done numerically (no closed-form expression); Use Table 4, Appendix III
  - Table is needed only on one side of the mean due to symmetry
  - The tabulated areas are to the right of points \( z \), where \( z \) is the distance from the mean, measured in standard deviations.

Example 4.8 Let \( Z \sim \text{N}(0, 1) \). Find
- \( a \ P(Z > 2) \)
- \( b \ P(-2 < Z < 2) \)
- \( c \ P(0 < Z < 1.73) \)

Example 4.9 The achievement score for a college entrance exam ~ \( \text{N}(75, 10^2) \) What fraction of the scores lies in \([80, 90]\)?
Recall that $z = (y - \mu)/\sigma$.

- Can always transform $Y \sim \text{N}(\mu, \sigma^2)$ to a “standard normal” RV $Z$ by using $Z = \frac{Y - \mu}{\sigma}$. Prove in Chapter 6 that the “standard normal RV” $Z \sim \text{N}(0, 1)$

- Exercises 4.46–66
- Keywords: $\text{N}(\mu, \sigma^2)$

### 4.6 The gamma probability distribution

- Some RVs are always
  - Nonnegative
  - Skewed (nonsymmetric) to the right
- E.g. lengths of time between malfunctions for aircraft engines; lengths of time between arrivals at a supermarket checkout queue
- Populations associated with these RVs frequently possess distributions that are adequately modeled by a gamma density function
- Definition 4.8 A RV $Y$ is said to have a “gamma distribution with parameters $a>0$ and $b>0$” or $Y \sim \text{Gamma}(a, b)$ iff
  \[ f(y) = \frac{1}{b^a \Gamma(a)} e^{-y/b} y^{a-1} I(y \geq 0) \] where
  \[ \Gamma(a) = \int_0^\infty e^{-y} y^{a-1} dy \] ("gamma function")
  - Integrations will verify that $\Gamma(1)=1$, $\Gamma(a)=(a-1) \Gamma(a-1)$ for any $a>0$, and $\Gamma(n)=(n-1)!$, for an integer $n$.
  - Parameter $a$: “shape parameter” (different graphs for $a=1, 2, 4$ and $b=1$)
  - Parameter $b$: “scale parameter”. Multiplying a gamma($a, b$) variable by say $c>0$, gives gamma($a, bc$) variable.

- Evaluating the integral
  - If $a$ is an integer, it is a sum of certain Poisson probabilities.
  - If $a$ is not an integer, no closed form. There’s a table but not in our book
- **Theorem 4.8** If $Y \sim \text{Gamma}(a, b)$, then $E(Y) = ab$ and $V(Y) = ab^2$
  - Proof: tricks a la Poisson/binomial mean/variance.
- **Definition 4.9** A random variable $Y$ is said to have a “chi-squared distribution with $v$ degrees of freedom” or $Y \sim \chi^2(v)$ iff $Y \sim \text{Gamma}(v/2, 2)$
- **Theorem 4.9** If $Y \sim \chi^2(v)$, then $E(Y) = v$ and $V(Y) = 2v$
- Use Table 6, Appendix III : can use this relationship to evaluate probability for SOME gamma distributions
Definition 4.10 A random variable $Y$ is said to have an "exponential distribution with parameter $b>0$" or $Y \sim \exp(\beta)$ iff $Y \sim \text{Gamma}(1, \beta)$ or $f(y) = \frac{1}{\beta} e^{-y/\beta} 1(y > 0)$.

- Useful for modeling the length of life of electronic components, especially when the length of a time a component already has operated does not affect its chance of operating for at least $b$ additional time units. "Memoryless" property.
- The geometric distribution also has the memoryless property. There's an interesting relationship between the two.

Theorem 4.10 If $Y \sim \exp(\beta)$, then $\mu = \mathbb{E}(Y) = \beta$ and $\sigma^2 = \mathbb{V}(Y) = \beta^2$.

Example 4.10 Let $Y \sim \exp(\beta)$. Show that if $a>0$ and $b>0$, $P(Y>a+b|Y>a)=P(Y>b)$.

Exercises 4.67~90

Keywords: $\exp(\beta)$

4.7 The beta probability distribution

- A two-parameter density function defined over the closed interval 0≤$y$≤1.
- Used to model proportions (examples?)

Definition 4.11 A random variable $Y$ is said to have a "beta probability distribution with parameters $\alpha>0$ and $\beta>0$" iff the density function of $Y$ is

$$f(y) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1}, 0 \leq y \leq 1 \text{ where } B(\alpha, \beta) = \int_0^1 y^{\alpha-1} (1-y)^{\beta-1} dy = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

- Shapes??
- Can obtain $y^*$ with range (0,1), and thus can be modeled as having beta distribution, from any continuous $y$ with range $(c,d)$ by $y^* = (y-c)/(d-c)$.
- The cumulative distribution function ("incomplete beta function") is

$$F(y) = \int_0^y t^{\alpha-1} (1-t)^{\beta-1} \frac{1}{B(\alpha, \beta)} = \sum_{i=\alpha}^{\alpha+\beta-1} \binom{n}{i} y^i (1-y)^{n-i} \text{ for } n=\alpha+\beta+1.$$ 

Theorem 4.11 If $Y \sim \text{Beta}(\alpha, \beta)$, then $\mu = \mathbb{E}(Y) = \frac{\alpha}{\alpha + \beta}$ and $\sigma^2 = \mathbb{V}(Y) = \frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$.

Example 4.11 The proportion of the gasoline supply that is sold during the week $\sim \text{Beta}(4,2)$.

P(selling at least 90% of the stock in a given week)=?

Exercises 4.91~103

Keywords: beta distribution; proportions
4.8 Some general comments

♦ Which theoretical distribution to use for a data? Rules of thumbs:
  □ Act on theoretical considerations
  □ Form a frequency histogram of the data

♦ Non-subjective statistical model selection procedures are available to test goodness of fit types hypotheses

4.9 Other expected values

♦ Moments (about the mean, about the origin) and the moment generating function (MGF) for continuous RVs have the same definitions as in the discrete case

♦ Example 4.13, 14 Find the MGF for $Y \sim \text{gamma}(a,b)$; also expand the MGF into a power series in $t$ and thereby obtain $\mu_k$.

♦ Example 4.15 The kinetic energy associated with a mass $m$ moving at velocity $v$ is given by $k = \frac{mv^2}{2}$. A device fires a serrated nail into concrete at a mean velocity of 2,000 ft/sec where the random velocity $V$ has a density function $f(v) = \frac{v^3 e^{-v/500}}{(500)^4 \Gamma(4)}$, $v > 0$. Find the expected kinetic energy associated with a nail of mass $m$.

♦ Theorem 4.12 The MGF for $g(Y)$ is $E[e^{g(Y)}] = \int_{-\infty}^{\infty} e^{g(y)} f(y) dy$.

♦ Example 4.16 Let $g(Y) = Y - \mu$ where $Y \sim N(\mu, \sigma^2)$. Find the MGF for $g(Y)$.

♦ Exercises 4.104~113

♦ Keywords: MGF

4.10 Tchebyshew’s theorem

♦ Theorem 4.13 Let $Y$ be a random variable with mean $\mu$ and finite variance $\sigma^2$. Then for any constant $k > 0$, $P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$ or $P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}$

  □ Proof:

♦ Example. Length of time (in minutes) required for a maintenance $Y \sim \text{gamma}(3.1, 2)$. A new worker takes 22.5 min for the task. Does this length of time disagree with prior experience?

  □ Or, evaluate $P(Y > 22.5)$

♦ HW. Some of the exercises 3.114~122.