Chapter 4

4.1 - Probability Density Function

Definition

A rv X is continuous if its set of possible values forms an interval or a union of intervals on a number line and the $P(X = \text{a number}) = 0$

Definition

The probability distribution (also known as the probability density function, pdf) of a continuous rv X is the function that satisfies the following

1. $f(x) \geq 0$ for any x
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. For any real numbers a and b with $a \leq b$

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

$$P(X = a) = \int_a^a f(x)dx = 0$$

$$P(a \leq X \leq b) = P(a < X < b)$$

Uniform Distribution

Definition

A rv X has a uniform distribution with parameters A and B if the pdf of X is

$$f(x; A, B) = \{^{\frac{1}{B-A},\ A \leq X \leq B}_{0,\ otherwise}\}$$

$$X \sim U[A, B]$$

Example 2

Given: $X =$ the reaction temperature

$$X \sim U[-5, 5]$$

1. $P(X < 0)$
   - $\int_{-5}^{0} \frac{1}{10} dx = 1/2$
2. $P(-2.5 \leq X \leq 2.5)$
   - $\int_{-2.5}^{2} .5 \frac{1}{10} dx = .5$
3. $P(-2 \leq X \leq 3)$
   - $\int_{-2}^{3} \frac{1}{10} dx = .5$

Example 6

The actual tracking weight of a stereo cartridge that is set to track at 3g on a particular changer can be regarded as a continuous rv X with pdf

$$f(x) = \{\substack{k[1-(x-3)^2] \quad 2 \leq x \leq 4 \\ 0 \quad otherwise}\}$$

1. Sketch the graph of $f(x)$

$$f(x) = \{\substack{\frac{3}{4}(1-(x-3)^2) \quad 2 \leq x \leq 4 \\ 0 \quad otherwise}\}$$

2. Find the value of k
   - $\int_2^4 k(1-(x-3)^2)dx = 1$
   - $k\int_2^4 (1-(x-3)^2)dx = 1$
   - $k(x|_2^4 - \frac{(x-3)^3}{3}|_2^4) = 1$
   - $k(2 - (\frac{1}{3} - \frac{(-1)^3}{3})) = 1$
   - $k(2 - \frac{2}{3}) = 1$
   - $k * \frac{4}{3} = 1$
   - $k = 3/4$
3. What is the probability that the actual tracking weight is greater than the prescribed weight?
   - $P(X > 3g) = \frac{1}{2}$
4. What is the probability that the actual weight is within .25g of the prescribed weight?
   - $|x - 3| < .25$
   - $P(2.75 < X < 3.25) =$
   - $= 2\int_{2.75}^{3} \frac{3}{4}(1-(x-3)^2)dx$
   - $= \frac{3}{2}(.25) - \frac{(x-3)^3}{3}|_{2.75}^{3}$
   - $= \frac{3}{2}(.25 - (0 - \frac{(-.25)^3}{3}))$
   - $= .3670$
5. What is the probability that the actual weight differs from the prescribed weight by more than .5g?
   - $|X - 3| > .5$
   - $P(X < 2.5 \text{ or } X > 3.5) = 2\int_2^{2.5} \frac{3}{4}(1-(X-3)^2)dx$
   - $= \frac{3}{2}(x|_2^{2.5} - \frac{(X-3)^3}{3}|_2^{2.5})$

- $= .3125$

4.2 - Cumulative Distribution and Expected Values

Definition

The cumulative distribution function of a rv X with the pdf $f(x)$ is

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

for any real x

Uniform Distribution

Definition

$$X \sim U[A, B]$$

$$F(x; A, B) = \{^{\frac{1}{B-A},\ A \leq x \leq B}_{0,\ otherwise}\}$$

If $x < A,\ then\ F(x; A, B) = 0$
If $A \leq x \leq B,\ F(x; A, B) = \int_A^x f(t; A, B)dt$
If $x \geq B,\ then\ F(x; A, B) = \int_A^B f(t; A, B)dt = 1$

Example 20

1. Compute and sketch the cdf of Y
   - $y < 0 \Rightarrow F(y) = 0$
   - $0 \leq y < 5 \Rightarrow F(y) = \int_0^y \frac{1}{25}tdt = \frac{t^2}{50}|_0^y = \frac{y^2}{50}$
   - $5 \leq y < 10 \Rightarrow F(y) = \int_0^y f(t)dt$
   - $= \int_0^5 \frac{1}{25}tdt + \int_5^y (\frac{2}{5} - \frac{1}{25}t)dt$
   - $= \ldots = \frac{1}{2} + \frac{2}{5}y - 2 = \frac{-y^2}{50} + \frac{1}{2}$
   - $\frac{2}{5}y - \frac{-y^2}{50} - 1$
   - $y \geq 10 \Rightarrow F(y) = 1$

Proposition

If X is a rv with the pdf $f(x)$ and the cdf $F(x)$, then

$$F^1(x) = f(x)$$

at every x at which F^1 exists
Note: let $a,\ b$ be real numbers, $a < b$

$$P(a \leq X \leq b) = F(b) - F(a)$$

$$P(a < X) = 1 - F(a)$$

Percentiles

ⓘ **Definition**

let $0 < p < 1$
The $(100p)^{th}$ percentile of the distribution of a rv X with the pdf $f(x)$ and the cdf $F(x)$ is the number

$$\eta(p)$$

is the number such that

$$F(\eta(p)) = p = \int_{-\infty}^{\eta(p)} f(x)dx$$

The median $\tilde{\mu}$ is the 50^{th} percentile

$$\tilde{\mu} = \eta(.5)$$

Example 20

2. Obtain an expression for the $(100p)^{th}$ percentile
    - $0 < p < .5 \Rightarrow p = \frac{y^2}{50} \Rightarrow y = \sqrt{50p} = 5\sqrt{2p}$
    - $\eta(p) = 5\sqrt{2p}$

4.3 - Normal Distribution

ⓘ **Definition**

Standard Normal Distribution

ⓘ **Definition**

Critical Values

ⓘ **Definition**

A critical value Z_a is the value of a rv $Z \sim N(0,1)$ with $P(Z > z_a) = \alpha$

Example

$z_a = 100(1-\alpha)^{th}$ percentile of $N(0,1)$
Find $z_{.10}(\alpha = .10)$

- $z_{.10}$ =the 90^{th} percentile of $N(0,1)$
- Approach 1
  - Table $A3$
  - $.8997 \Rightarrow z_{.10} = 1.28$
- Approach 2
  - Table $A3$
  - $.1003 \Rightarrow -z_{.10} = -1.28 \Rightarrow z_{.10} = 1.28$

When multiple numbers on the table are equidistant from the number you need, you can average the z-scores

Table 4.1 $(p.161)$ provides common critical values

Standardizing a Random Variable

Proposition

If $X \sim N(\mu, \sigma^2)$, then

$$z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$P(a \leq X \leq b) = P(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}) = \phi(\frac{b-\mu}{\sigma}) - \phi(\frac{a-\mu}{\sigma})$$

$$P(X \leq b) = P(z \leq \frac{b-\mu}{\sigma}) = \phi(\frac{b-\mu}{\sigma})$$

Example 32

X = the force, $X \sim N(\mu = 15, \sigma^2 = 1.25^2)$

1. $P(X \leq 15)$
   - $\frac{1}{2}$
2. $P(X \leq 17.5)$
   - $P(Z \leq \frac{17.5-15}{1.25})$

- $= \phi(2) = .9772$

3. $P(X \geq 10)$
   - $= 1 - P(X < 10) = P(Z < \frac{10-15}{1.25} = 1 - \phi(-4)) \approx 1$
4. $P(14 \leq X \leq 18)$
   - $= P(\frac{14-15}{1.25} \leq Z \leq \frac{18-15}{1.25})$
   - $= P(-8 \leq Z \leq 2.4) = \phi(2.4) - \phi(-.8)$
   - $= .9918 - .2119 = .7799$
5. $P(|X - 15| \leq 3)$
   - $= P(12 \leq X \leq 18) = \phi(\frac{18-15}{1.25}) - \phi(\frac{12-15}{1.25})$
   - $= \phi(2.4) - \phi(-2.4)$

Establishing Connections between Percentiles

Definition

Let $Z \sim N(0, 1)$. Then

$$X = \mu + \sigma Z \sim N(\mu, \sigma)$$

$$[^{The\ (100p)^{th}\ percentile}_{of\ N(\mu, \sigma^2)}] = \mu + \sigma[^{The\ (100p)^{th}\ percentile}_{of\ N(0,1)}]$$

Example 40

Let X = the yield strength
$X \sim N(\mu = 43, \sigma^2 = 4.5^2)$
c = the 25^{th} percentile of $N(43, 4.5^2)$
Find the 25^{th} percentile of $N(0, 1)$

- Table $A3$ $(.2514) \implies z_{.75} = -.67$
- $c = 43 + 4.5(-.67) = 39.99$

Empirical Rule

Definition

If the population distribution of a variable is (approximately) normal, then

1. Roughly 68% of the values are within 1 SD of the mean
2. Roughly 95% of the values are within 2 SDs of the mean
3. Roughly 99.7% of the values are within 3 SDs of the mean

Continuity Correction

Proposition

Let X be a binomial rv based on n trials with success probability p. Then if the binomial probability histogram is not too skewed, X has approximately a normal distribution with $\mu\ldots$

Example

$P(x \leq 12) = B(12; 20, 0.5) \approx \phi(\frac{12+0.5-10}{2.236})$

Example 50

$n = 1000$
$S =$ a person can taste the difference
$p = P(S) = .03$
$q = 1 - p = .97$

1. What is the probability that at least 40 can taste the difference?
   - $P(X \geq 40)$ if $X \sim Bin(1000, .03)$
   - $P(X \geq 40) = 1 - P(X \leq 39) = 1 - B(39; 1000, .03)$
   - $\approx 1 - \phi(\frac{39+.5-30}{\sqrt{1000(.03)(.97)}}) = 1 - \phi(1.76) = 1 - .9608 = .0392$
2. What is the probability that at most 50 can taste the difference?
   - $P(X \leq 50) = B(50; 1000, .03)$
   - $\approx \phi(\frac{50+.5-30}{sqrt1000(.03)(.97)} = \phi(3.80) \approx 1$

4.4 - The Exponential and Gamma Distributions

Gamma Distributions

Exponential Distribution

Parameter: λ, $\lambda > 0$

The pdf:

$$f(x) = \{^{\lambda e^{-\lambda x},\ x \geq 0}_{0,\ x<0}\}$$

Notation:

$$X \sim exp(\lambda)$$

The cdf

$$x < 0,\ F(x;\lambda) = P(X \le x) = 0$$

$$x \ge 0,\ F(x;\lambda) = \int_0^x \lambda e^{-\lambda t} dt = \ldots = 1 - e^{-\lambda x}$$

$$F(x;\lambda) = \{{}^{1-e^{-2\lambda},\ x \ge 0}_{0,\ x<0}\}$$

Proposition

For $X \sim exp(\lambda)$

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda_2}$$

$$\sigma_X = \frac{1}{\lambda}$$

Example 60

$X \sim exp(\lambda = .01386)$

1. - $P(X \le 100) = F(100, \lambda)$
     - $= 1 - e^{-.01386(100)} = .750$
   - $P(X \le 200) = F(200, \lambda)$
     - $1 - e^{-.01386(200)} = .937$
   - $P(100 \le X \le 200)$
     - $= .937 - .750 = .187$
2. Probability that distance exceeds the mean distance by more than 2 standard deviations?
   - $E(x) = \frac{1}{\lambda} = \frac{1}{.01386} \approx 72.150m \approx \sigma$
   - $P(X > \mu + 2\sigma) = P(X > 3(72.150))$
   - $= P(X > 216.45) = 1 - F(216.45;\lambda) = \ldots = .050$
3. Find the median distance
   - $F(\tilde{\mu}) = .5 \Rightarrow 1 - e^{-.01386\tilde{\mu}} = .5$
   - $\Rightarrow e^{-.01386\tilde{\mu}} = .5 \Rightarrow -.01386\tilde{\mu} = ln.5$
   - $\tilde{\mu} = \frac{ln.5}{.01386} \approx 50.01\ meters$

Gamma Function

ⓘ Definition

The gamma function is

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$$

ⓘ Properties

1. For any $\alpha > 1,\ \Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
2. For any integer $n \geq 1,\ \Gamma(n) = (n-1)!$
3. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Examples

1. $\Gamma(6) = (6-1)! = 5! = 120$
2. $\Gamma(\frac{5}{2})$
   - Use rule 1 twice then rule 3

Gamma Distribution

Parameters: $\alpha, \beta > 0$

The pdf

$$f(x;\alpha,\beta) = \{ {\frac{1}{\beta\Gamma(\alpha)}x^{\alpha-1}e^{\frac{-x}{\beta}},\, x\geq 0 \atop 0, x<0} \}$$

Note: $\alpha \Rightarrow exp(\lambda = \frac{1}{\beta})$
Notations:

$$X \sim Gamma(\alpha;\beta)$$

$$X \sim \Gamma(\alpha,\beta)$$

The cdf

$$F(x;\alpha,\beta) = P(X \leq x)$$

The Mean and variance

$$E(X) = \alpha\beta$$

$$V(X) = \alpha\beta^2$$

$$\sigma_X = \alpha\beta$$

Standard Gamma Distribution

$\Gamma(\alpha,\ \beta = 1)$

The pdf

$$x \geq 0 : f(x; \alpha) = \frac{1}{\Gamma(\alpha)} x^{\alpha - 1} e^{-x}$$

The cdf

Known as the incomplete gamma function

$$F(x; \alpha) = P(X \leq x) = \int_0^x \frac{1}{\Gamma(\alpha)} t^{\alpha - 1} e^{-t} dt$$

Proposition

Let $X \sim Gamma(\alpha, \beta)$
Then

$$P(X \leq x) = F(x; \alpha, \beta)$$
$$= F(\frac{x}{\beta}; \alpha)$$

This is the incomplete gamma function
Consult Table $A4$

Example 67

$X =$ the lifetime in weeks
$X \sim Gamma(\alpha,\ \beta)$
$E(X) = 24$
$\sigma = 12$

1. $P(12 \leq X \leq 24)$
    - $24 = \alpha\beta$
    - $144 = \alpha\beta^2$
    - $\alpha = 4,\ \beta = 6$
    - $P(12 \leq X \leq 24)$
        - $= F(\frac{24}{\sigma}; \alpha = 4) - F(\frac{12}{\sigma}; \alpha = 4)$
        - $= F(4; 4) - F(2; 4) = .567 - .143 = .424$
2. - $P(X \leq 24)$
        - $F(\frac{24}{\sigma}; \alpha = 4) = .567$
    - Is $\tilde{\mu} < 24$?

- $F(\frac{\tilde{\mu}}{\sigma}; \alpha = 4)$
- $= F(\tilde{\mu}; \alpha = 4,\ \beta = 6) = .5$
- Yes

3. - Find $\eta(.99)$
    - $F(\eta(.99), \alpha = 4, \beta = 6) = .99$
    - $F(x; \alpha = 4, \beta = 6) = F(\frac{x}{6]}, \alpha = 4) = .99$
    - Table $A4 \Rightarrow F(10, \alpha = 4) = .99$
    - $\frac{x}{6} = 10 \Rightarrow x =$ the 99^{th} percentile = 60 weeks
    - 99% of transistors subjected to an accelerated life tests have lifetimes that are less than or equal to 60 weeks

4. - Find t such that $P(X > t) = .005$
   - $P(X \leq t = .995$
   - $F(t; 4, 6) = F(\frac{t}{6}, \alpha = 4) = .995$
   - Table $A4 \Rightarrow \frac{t}{6} = 11 \Rightarrow t = 66$ weeks