Chapter 4: Continuous Variables and Their Probability Distributions

Chapter 4 extends probability theory from discrete to continuous random variables, introducing the mathematical machinery necessary to model measurements and quantities that can take any value within an interval.

Introduction and Fundamental Concepts (Sections 4.1-4.2)

The transition from discrete to continuous random variables requires replacing probability mass functions with probability density functions (PDFs). For continuous variables, individual point probabilities are zero; instead, we work with probabilities over intervals. The probability density function $f(y)$ satisfies $f(y) \geq 0$ and $\int_{-\infty}^{\infty} f(y)dy = 1$. Probabilities are computed as areas under the density curve: $P(a \leq Y \leq b) = \int_a^b f(y)dy$.

The cumulative distribution function (CDF) $F(y) = P(Y \leq y) = \int_{-\infty}^{y} f(t)dt$ provides an alternative representation. A key property distinguishing continuous from discrete variables: for continuous random variables, $P(Y = c) = 0$ for any specific value c, which means $P(a < Y < b) = P(a \leq Y \leq b)$—endpoint inclusion doesn't affect probability.

Expected Values (Section 4.3)

Expected values extend continuous variables, replacing summation with integration. For a continuous random variable Y with density $f(y)$:

- Mean: $E(Y) = \mu = \int_{-\infty}^{\infty} yf(y)dy$
- Variance: $V(Y) = \sigma^2 = \int_{-\infty}^{\infty}(y-\mu)^2 f(y)dy = E(Y^2) - [E(Y)]^2$
- General expectations: $E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy$

Linear transformations preserve the linearity properties: $E(aY+b) = aE(Y)+b$ and $V(aY+b) = a^2V(Y)$.

The Uniform Distribution (Section 4.4)

The uniform distribution represents complete uncertainty over an interval (θ_1, θ_2), with constant density:

$$f(y) = \frac{1}{\theta_2 - \theta_1}, \quad \theta_1 \leq y \leq \theta_2$$

Key properties include $E(Y) = \frac{\theta_1+\theta_2}{2}$ and $V(Y) = \frac{(\theta_2-\theta_1)^2}{12}$.

The Normal Distribution (Section 4.5)

The normal (Gaussian) distribution is

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}, \quad -\infty < y < \infty$$

The standard normal distribution has $\mu = 0$ and $\sigma = 1$. Any normal variable can be standardized via $Z = \frac{Y-\mu}{\sigma}$, enabling probability calculations using standard normal tables. The normal distribution is symmetric about its mean, with approximately 68% of observations within one standard deviation, 95% within two standard deviations, and 99.7% within three standard deviations (the empirical rule).

The Gamma Distribution (Section 4.6)

The gamma distribution models waiting times and durations, with density:

$$f(y) = \frac{y^{\alpha-1} e^{-y/\beta}}{\beta^\alpha \Gamma(\alpha)}, \quad y > 0$$

where $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$ is the gamma function. Parameters are α (shape) and β (scale), with $E(Y) = \alpha\beta$ and $V(Y) = \alpha\beta^2$.

Special cases include:

- Exponential distribution: $\alpha = 1$, modeling time between events
- Chi-squared distribution: $\alpha = \nu/2$, $\beta = 2$, fundamental in statistical inference

The exponential distribution is memoryless: $P(Y > s + t | Y > s) = P(Y > t)$.

The Beta Distribution (Section 4.7)

The beta distribution is defined on the interval $(0, 1)$, used for proportional probabilities

$$f(y) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 < y < 1$$

where $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ is the beta function. The mean is $E(Y) = \frac{\alpha}{\alpha+\beta}$ and variance is $V(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

Note: symmetric when $\alpha = \beta$, skewed right when $\alpha < \beta$, skewed left when $\alpha > \beta$. The special case $\alpha = \beta = 1$ reduces to the uniform distribution.

Distribution Relationships (Section 4.8)

Several distributions are interconnected through limiting processes or parameter specializations. The normal approximates the binomial for large n. The exponential is a special case of the gamma. Linear transformations of normal variables remain normal.

Moment-Generating Functions (Section 4.9)

Moment-generating functions $m(t) = E(e^{tY})$ Key MGFs include:

- Exponential: $m(t) = (1 - \beta t)^{-1}$
- Normal: $m(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$
- Gamma: $m(t) = (1 - \beta t)^{-\alpha}$

MGFs uniquely determine distributions and facilitate proving that linear combinations of independent normal variables are normal.

Tchebysheff's Theorem (Section 4.10)

Tchebysheff's theorem provides distribution-free probability bounds:

$$P(|Y - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

This applies to any distribution with finite variance, guaranteeing that at least 75% of observations lie within two standard deviations of the mean, and at least 89% within three standard deviations. While conservative for specific distributions (like the normal), it provides useful bounds when the exact distribution is unknown.

Chapter 4 in a Nutshell

Integrate where you would sum in discrete probability distributions.