February 25, 27

Instructor: Brian Powers

15.1 Continuous Probability Distributions

15.1.1 Uniform Distribution

Definition 15.1. A random variable X which can take any real number value from a to b with equal probability is said to follow a **continuous uniform distribution**. As notation we say $X \sim Unif_C(a, b)$.

Theorem 15.2. *If $X \sim Unif_C(a, b)$, its pdf is given as*

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \textit{elsewhere} \end{cases}$$

Its expected value and variance are

$$E(X) = \frac{a+b}{2}, Var(X) = \frac{(b-a)^2}{12}.$$

Proof. The pdf can be found by assuming $f(x) = c$ on the support $[a, b]$, and integrating to find c.

$$E(X) = \int_a^b x\frac{1}{b-a} = \frac{1}{b-a}\left[\frac{1}{2}x^2\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}$$

The proof for variance is omitted, but it is proved in the same way. □

15.1.2 Normal Distribution

Definition 15.3. A random variable X which can take any real value with probability density according to the well known "bell curve" follows a **Normal distribution**. The parameters of the distribution are its mean μ and its variance σ^2. To say X follows such a normal distribution we denote this as $X \sim N(\mu, \sigma^2)$. The pdf is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

The normal distribution is completely determined by μ (its location) and σ (its scale).

Theorem 15.4. *If $X \sim N(\mu, \sigma^2)$,*

$$E(X) = \mu, \quad Var(X) = \sigma^2.$$

The proof is straightforward calculus.

- The mean, median and mode of $N(\mu, \sigma^2)$ is μ.

- The curve of the pdf is symmetric about $x = \mu$.

Definition 15.5. A normal random variable is said to follow the **standard normal distribution** if $\mu = 0$ and $\sigma^2 = 1$, and we traditionally denote such a random variable as Z, thus, $Z \sim N(0, 1)$. The pdf of the standard normal distribution is

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

Standardizing means to center around zero then scale so that the standard deviation (and variance) is 1. Thus, for $X \sim N(\mu, \sigma^2)$, the standardized random variable is

$$Z = \frac{X - \mu}{\sigma}.$$

For any value x that X may take, the corresponding standardized value of Z is $z = (x - \mu)/\sigma$.

Theorem 15.6. *For $X \sim N(\mu, \sigma^2)$, and $x_1 < x_2 \in \mathbb{R}$,*

$$P(x_1 < X < x_2) = P(z_1 < Z < z_2),$$

where Z follows a standard normal distribution, and $z_i = (x_i - \mu)/\sigma$.

Proof.

$$\begin{aligned} P(x_1 < X < x_2) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{x_1}^{x_2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \\ &= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{1}{2}z^2} \qquad \text{Change of variables } z = (x-\mu)/\sigma, dz = dx/\sigma \\ &= P(z_1 < Z < z_2) \end{aligned}$$

□

Because any normal variable can be standardized for the purpose of evaluating probabilities, we will consider only the cdf of the standard normal distribution. Let

$$F_Z(z) = \Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \phi(t)dt$$

The integral has no closed form in general, instead we use technology or tables to give us the evaluations.

Example 15.7. Derive the "Empirical Rule" of 68%, 95%, 99%.

As notation, we will let z_α be the α quantile of the standard normal distribution. In other words,

$$P(Z < z_\alpha) = \alpha, \quad \text{or } z_\alpha = \Phi^{-1}(\alpha).$$

If we wish to find the corresponding quantile for the distribution of $X \sim N(\mu, \sigma^2)$, we use

$$x_\alpha = \sigma z_\alpha + \mu.$$

15.1.3 Gamma Distribution

Definition 15.8. The **gamma function** is defined as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \quad \alpha > 0.$$

It has the following properties:

- Recursive: $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
- Integers: $\Gamma(n) = (n-1)!$ for an integer $n > 0$.
- $\Gamma(1/2) = \sqrt{\pi}$.

Definition 15.9. The **gamma distribution** takes two parameters, α, and β. If $X \sim \mathit{Gamma}(\alpha, \beta)$, its pdf is

$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x > 0 \\ 0 & \text{elsewhere,} \end{cases}$$

Where $\alpha > 0, \beta > 0$.

The gamma distribution is useful to model X, the length of time until k occurrences of a rare event under similar assumptions as the Poisson process where λ is the mean number of occurrences during 1 unit of time. $X \sim \mathit{Gamma}(\alpha = k, \beta = 1/\lambda)$.

Theorem 15.10. *If $X \sim \mathit{Gamma}(\alpha, \beta)$,*

$$E(X) = \alpha\beta, \quad Var(X) = \alpha\beta^2.$$

The proof is beyond this course and more suited to a higher level course in statistics. For integer values of α, Gamma distribution cdf probabilities may be evaluated. The regularized lower **incomplete gamma function** is

$$F(x; \alpha) = \int_0^x \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)} dy.$$

Example 15.11. Say $X \sim \mathit{Gamma}(5, 10)$. Find $P(X \le 60)$.

$$\begin{aligned} P(X \le 60) &= \int_0^{60} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx \\ &= \int_0^6 \beta \frac{(x/\beta)^{\alpha-1} e^{-(x/\beta)}}{\Gamma(\alpha)} \frac{dx}{\beta} \\ &= \int_0^6 \frac{y^4 e^{-y}}{\Gamma(5)} dy \qquad \text{substitute } y = x/\beta, dy = dx/\beta \\ &= F(6; 5) \\ &\approx 0.715 \end{aligned}$$

We can use tables for the incomplete gamma function. It is related to the poisson cdf. For integer values α,

$$F(x; \alpha) = 1 - \texttt{poissoncdf}(\texttt{x}, \alpha - \texttt{1})$$

15.1.4 Exponential Distribution

Definition 15.12. The **exponential distribution** is a special case of the gamma distribution where $\alpha = 1$. Let $X \sim exponential(\beta)$. Then pdf is given by

$$f(x) = \begin{cases} \frac{1}{\beta}e^{-x/\beta} & x > 0 \\ 0 & \text{elsewhere,} \end{cases}$$

Where $\beta > 0$.

The exponential distribution models the time between occurences in a Poisson process with $\lambda = 1/\beta$.

Corollary 15.13. *If* $X \sim exponential(\beta)$*,*

$$E(X) = \beta, Var(X) = \beta^2.$$

15.1.5 Chi-Squared Distribution

Definition 15.14. The **chi-squared distribution** is a special case of the gamma distribution with $\alpha = v/2, \beta = 2$. We say that X follows a chi-squared distribution with v degrees of freedom, denoted $X \sim \chi^2(v)$.

Corollary 15.15. *If* $X \sim \chi^2(v)$*,*

$$E(X) = v, Var(X) = 2v.$$