Tuesday, January 22: Distribution families

Course topics:

1. Distribution families
2. Interval estimates
3. Point estimates
4. Hypothesis tests
5. Linear regression

By the end of the course, we will have explored the mathematics behind the material in Math 160 and Math 260
Tuesday, January 22: Distribution families

You should be familiar with the following terms regarding random processes:

- random process
- outcome
- event
- sample space
- probability function
- conditional probability
- independent
Tuesday, January 22: Distribution families

And the following terms regarding random variables:

- random variable
- discrete and continuous
- distribution
- probability mass function and probability density function
- cumulative distribution function
- quantiles
- conditional distribution
- independent
- expected value (or mean)
- variance and standard deviation
- moments
Some distribution families to know

The **uniform** distribution family $\text{Unif}(a, b)$ is a 2-parameter family of continuous distributions, where $a, b$ are real numbers satisfying $b > a$.

A continuous random variable $X$ is **uniformly distributed** on $[a, b]$ (written $X \sim \text{Unif}(a, b)$) if its probability density function is

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{otherwise.} \end{cases}$$
Tuesday, January 22: Distribution families

In R, the suffix for a uniform distribution is `unif`, and its parameters are `min` and `max`.

This means that R has the following functions:

- `dunif(x, min=a, max=b)` gives the value of the probability density function of $\text{Unif}(a, b)$ at $x$.
- `punif(x, min=a, max=b)` gives the value of the cumulative distribution function of $\text{Unif}(a, b)$ at $x$.
- `qunif(p, min=a, max=b)` gives the $p$-th quantile of $\text{Unif}(a, b)$.
- `runif(n, min=a, max=b)` draws a (pseudo-random) sample of size $n$ from a $\text{Unif}(a, b)$-distribution.

More information can be obtained from the help page, accessed with `?punif` (or any of the other functions).
Tuesday, January 22: Distribution families

The **binomial** distribution family $B(n, p)$ is a 2-parameter family of discrete distributions, where the **size** $n$ is a positive integer and the **success probability** $p$ is a real number with $0 < p < 1$.

A discrete random variable $X$ is **binomially distributed** with size $n$ and success probability $p$ (written $X \sim B(n, p)$) if its probability mass function is

$$f(k) = \begin{cases} \binom{n}{k}p^k(1 - p)^{n-k} & \text{if } k \in \{0, 1, \ldots, n\} \\ 0 & \text{otherwise.} \end{cases}$$

In R, the suffix for a binomial distribution is `binom`, and its parameters are `size` and `prob`
The following two propositions are exercises for the reader:

**Proposition** If \( X \) is the number of successes in a sequence of \( n \) independent Bernoulli trials each with success probability \( p \), then \( X \sim B(n, p) \).

**Proposition** If \( X_1, \ldots, X_n \) are independent and identically distributed discrete random variables each with probability mass function

\[
f(k) = \begin{cases} 
1 - p & \text{if } k = 0 \\
p & \text{if } k = 1 \\
0 & \text{otherwise},
\end{cases}
\]

then

\[
\sum_{i=1}^{n} X_i \sim B(n, p).
\]
The multinomial distribution family $M(n, p_1, p_2, \ldots, p_r)$ is an $(r + 1)$-parameter family of discrete joint distributions, where the size $n$ is a positive integer and the success probabilities $p_1, \ldots, p_r$ are real numbers with $0 < p < 1$ satisfying $\sum_{i=1}^{r} p_i = 1$.

A set of discrete random variables $X_1, \ldots, X_r$ are jointly multinomially distributed with size $n$ and success probabilities $p_1, \ldots, p_r$ (written $X \sim M(n, p_1, \ldots, p_r)$) if their joint probability mass function is

$$f(k_1, \ldots, k_r) = \begin{cases} \frac{n!}{k_1! \cdots k_r!} p_1^{k_1} \cdots p_r^{k_r} & \text{if } k_1, \ldots, k_r \in \{0, 1, \ldots, n\} \\ 0 & \text{with } \sum_{i=1}^{r} k_i = n \text{ otherwise.} \end{cases}$$
In R, the suffix for a multinomial joint distribution is `multinom`, and its parameters are `size` and `prob` (which is a vector of probabilities, which will be automatically normalized to sum to 1).

Since this is a joint distribution, there are no `pmultinom` or `qmultinom` functions.

We have the following proposition:

**Proposition** Suppose that a random process has exactly $r$ types of outcomes, which occur with probabilities $p_1, \ldots, p_r$. For each $i \in \{1, 2, \ldots, r\}$, let $X_i$ be the number of outcomes of type $i$ in a sequence of $n$ independent trials of this random process. Then the joint distribution of $X_1, \ldots, X_r$ is $X \sim M(n, p_1, \ldots, p_r)$. 
The **Poisson** distribution family Poisson($\lambda$) is a 1-parameter family of discrete distributions, where the **mean** $\lambda$ is a positive real number.

A discrete random variable $X$ is **Poisson distributed** with mean $\lambda$ (written $X \sim \text{Poisson}(\lambda)$) if its probability mass function is

$$f(k) = \begin{cases} \frac{\lambda^k e^{-\lambda}}{k!} & \text{if } k \in \mathbb{Z} \text{ and } k \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Poisson distributions often arise when counting occurrences per time interval of an event that has, roughly speaking, small but constant probability over a continuous time.

In R, the suffix for a Poisson distribution is `pois`, and its parameter is `lambda`
The **gamma function** $\Gamma : \mathbb{R} \to \mathbb{R}$ is defined by:

$$\Gamma(x) = \int_0^\infty y^{x-1} e^{-y} \, dy.$$ 

In the homework, you will explore some of the basic properties of the gamma function, such as:

$$\Gamma(n) = (n-1)!$$

for all positive integers $n$, and

$$\Gamma(1/2) = \sqrt{\pi}.$$ 

In R, you can get approximate values of the gamma function with `gamma()`.
The **gamma** distribution family $\text{Gamma}(\alpha, \beta)$ is a 2-parameter family of continuous distributions, where the **shape parameter** $\alpha$ is a positive real number and the **rate parameter** $\beta$ is a positive real number.

A continuous random variable $X$ is **gamma distributed** with shape parameter $\alpha$ and rate parameter $\beta$ (written $X \sim \text{Gamma}(\alpha, \beta)$) if its probability density function is

$$f(x) = \begin{cases} 
\frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & \text{for } x > 0 \\
0 & \text{otherwise.}
\end{cases}$$
In R, the suffix for a gamma distribution is `gamma`, and its parameters are `shape` and `rate`.

Gamma distributions form a broad family of distributions that includes some other familiar distribution families.
Friday, January 25: Distribution families

The **exponential** distribution family $\text{Expon}(\theta)$ is a 1-parameter family of continuous distributions, where the **mean** $\theta$ is a positive real number.

A continuous random variable $X$ is **exponentially distributed** with mean $\theta$ (written $X \sim \text{Expon}(\theta)$) if its probability density function is

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & \text{for } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Exponential distributions are specific types of gamma distributions:

$$\text{Expon}(\theta) = \text{Gamma}(1, \theta).$$
In R, the suffix for an exponential distribution is `exp`, and its parameter is `rate`.

Exponential distributions are often used to model times until events and times between events in a Poisson process.
Friday, January 25: Distribution families

The chi-square distribution family $\chi^2(k)$ is a 1-parameter family of continuous distributions, where the degrees of freedom $k$ is a positive real number.

A continuous random variable $X$ is chi-square distributed with $k$ degrees of freedom (written $X \sim \chi^2(k)$) if its probability density function is

$$f(x) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)}x^{(k-2)/2}e^{-x/2} & \text{for } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Chi-square distributions are specific types of gamma distributions:

$$\chi^2(k) = \text{Gamma}(k/2, 2).$$
Friday, January 25: Distribution families

In R, the suffix for a chi-square distribution is `chisq`, and its parameter is `df`.

We will see later that chi-square distributions are closely related to normal distributions.