Introduction to Probability

Traditional Statistics Review

- Probability.
- Z-test, t-test, and analysis of variance (ANOVA).
- Regression.

Prerequisites

- Some mathematics are required.
- For example,

$$\int x dx = \frac{1}{2}x^2.$$

$$\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}} \quad (a > 0).$$

$$\int_0^{\infty} \frac{1}{x} dx = \infty.$$

Prerequisites

- Normal distribution, $N(\mu, \sigma^2)$:

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

- Beta distribution, $Beta(\alpha, \beta)$:

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad x \in (0,1)$$

- Gamma distribution, $Gamma(\alpha, \beta)$:

$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad x \in (0, \infty)$$

Prerequisites: Probability

- Probabilities apply to processes with **unpredictable** outcomes ("random experiments")
- Probability model:
  1. Random variable X (the result, or outcome).
  2. Sample Space $\mathcal{X}$ (Set of all possible outcomes).
  3. Probability distribution over $\mathcal{X}$.
- e.g.: Coin flipping.

Prerequisites: Probability

- The probability of an event (set) A, $P(A)$, is the sum of probabilities of all the points that are in A.
- e.g.: dice rolling.

Prerequisites: Probability

- Suppose we select one student at random from those registered for this class and determine the number of teeth in that person's head. The result of this process will be a number. Let's call it X. This is our **Random Variable**. Consider the sample space is $\mathcal{X} = \{0, 1, 2, ..., 30, 31, 32\}$. Let $P(X = 0)$ be the proportion of students with no teeth, $P(X = 1)$ be the proportion of students with one tooth and so on.
- The event "selected student has at least 26 teeth" is represented by the set

$$A = \{26, 27, 28, 29, 30, 31, 32\}.$$

Prerequisites: Probability

- The event "Selected student has at least 26 teeth or has an even number of teeth" is represented by the set:

$$A \text{ or } B = \{26, 27, 28, 29, 30, 31, 32, 0, 2, ..., 22, 24\}.$$

- Its probability is

$$\begin{aligned} & P(A \cup B) = P(A) + P(B) - P(A \cap B) \\ = \quad & [P(26) + P(27) + ... + P(32)] + [P(0) + P(2) + ... + P(32)] \\ & \quad -[P(26) + P(28) + P(30) + P(32)] \end{aligned}$$

Prerequisites: Probability

Properties of probabilities:

- For event A in sample space $\mathcal{X}$,
  1. $0 \leq P(A) \leq 1$.
  2. $P(\mathcal{X}) = 1$.
  3. $P(A) = 1 - P(A^c)$.
  4. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

Prerequisites: Probability

Properties of probabilities

- For events A and B,
  1. If $P(A \text{ and } B) = P(A)P(B)$, then A and B are independent events.
  2. If $P(A \text{ and } B) = 0$, then A and B are mutually exclusive or disjoint.
  3. The conditional probability of A given B is

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}.$$

Example

- A woman and a man (unrelated) each have two children. At least one of the woman's children is a boy, and the man's older child is a boy. Do the chances that the woman has two boys equal the chances that the man has two boys?
- Marilyn says: The chances that the woman has two boys are 1 in 3 and the chances that the man has two boys are 1 in 2.
- Many people write in to tell Marilyn that she is horribly wrong and a disgrace to the human race. Obviously the chances are equal. Who is correct?

Example

- **Assumptions**: For any family, the probability of a boy on one birth is $\frac{1}{2}$, and births are independent.
- **Notations**:
  1. $\mathcal{X} = \{(0,0),(0,1),(1,0),(1,1)\}$
  2. Events:

$$A = \{\text{older birth is a boy}\} = \{(0,1),(1,1)\}$$

$$C = \{\text{Exactly one boy in two births}\} = \{(1,0),(0,1)\}$$

$$D = \{\text{Exactly two boys in two births}\} = \{(1,1)\}.$$

Example

- We are given that the man's older child is a boy. What is the probability of two boys, given the older is a boy?
- Mathematically, it is $P(D \mid A)$.
- Answer:

$$\begin{aligned} P(D \mid A) &= \frac{P(D \cap A)}{P(A)} \\ &= \frac{P(D)}{P(A)} = \frac{1/4}{1/2} = \frac{1}{2}. \end{aligned}$$

Example

- We are also given that the woman has at least one boy. What is the probability of two boys, given at least one boy?
- Mathematically, it is $P(D \mid C \cup A)$.
- Answer:

$$\begin{aligned} P(D \mid C \cup A) &= \frac{P(D \cap \{C \cup A\})}{P(C \cup A)} \\ &= \frac{P(D)}{P(C \cup A)} = \frac{1/4}{3/4} = \frac{1}{3}. \end{aligned}$$

- Marilyn is Correct!

Law of Total

- Suppose the sample space is divided into any number of disjoint sets, say $A_1, A_2, ..., A_n$, so that $A_i \cap A_j = \emptyset$ and $A_1 \cup A_2 \cup ... \cup A_n = \mathcal{X}$.
- In this case we can write

$$P(B) = \sum_{i}^{n} P(B \cap A_i)$$

- In other words,

$$P(B) = \sum_{i} P(B \mid A_i) P(A_i)$$

Example: Law of Total

- Suppose that two factories supply light bulbs to the market. Factory X's bulbs work for over 5000 hours in 99% of cases, whereas factory Y's bulbs work for over 5000 hours in 95% of cases. It is known that factory X supplies 60% of the total bulbs available. What is the chance that a purchased bulb will work for longer than 5000 hours?

Example: Law of Total

- Applying the law of total probability, we have:

$$\begin{aligned}\Pr(A) &= \Pr(A \mid B_X) \cdot \Pr(B_X) + \Pr(A \mid B_Y) \cdot \Pr(B_Y) \\ &= \frac{99}{100} \cdot \frac{6}{10} + \frac{95}{100} \cdot \frac{4}{10} = \frac{594 + 380}{1000} = \frac{974}{1000}\end{aligned}$$

- Thus each purchased light bulb has a 97.4% chance to work for more than 5000 hours.

Estimation for Population Proportion

- The population proportion: $P = \frac{X}{N}$ where X is the count of successes in the population and N is the size of the population.
- The sample proportion: $\hat{p} = \frac{x}{n}$ where x is the count of successes in the sample and n is the size of the sample obtained from the population.

Example: Estimation for Population Proportion

- Suppose that 6 out of 40 students plan to go to graduate school. Then what would be the proportion of all students who plan to go to graduate school?
- What would be the standard deviation for the estimation?
- What would be the 95% Confidence Interval for the estimation?

Example: Estimation for Population Proportion

- Suppose a presidential election is taking place in a democracy. A random sample of 400 eligible voters in the democracy's voter population shows that 272 voters support candidate B. A political scientist wants to determine what percentage of the voter population support candidate B.
- What would be the 95% Confidence Interval for the estimation?

Statistical Hypothesis Testing

- We usually do not know the true value of population parameters - they must be estimated. However, we do have hypotheses about what the true values are.
- We do this by calculating the **p-value**, the probability of the **data** if the null hypothesis is true.

Statistical Hypothesis Testing

- The **p-value** is the probability that a given result (or a more significant result) would occur under the null hypothesis.
- For example, say that a fair coin is tested for fairness (H_0). At a significance level of 0.05, the fair coin would be expected to (incorrectly) reject the null hypothesis in about 1 out of every 20 tests.
- The p-value does not provide the probability that either hypothesis is correct.

Example: Baby Birth Weight

- From previous experience we know that the birth weights of babies in England are Normally distributed with a mean of 3000g and a standard deviation of 500g.
- We think that maybe babies in Australia have a mean birth weight greater than 3000g and we would like to test this hypothesis.

Example: Baby Birth Weight

- Setting up the hypotheses:

$$H_0 : \mu = 3000g$$
$$H_1 : \mu > 3000g$$

- Suppose that we take a sample of 44 babies from Australia, measure their birth weights and we observe that the sample mean of these 44 weights is

$$\bar{X} = 3275.955g.$$

Example: Baby Birth Weight

- Under the null hypothesis, the sample mean of 44 values from a $N(3000, 500^2$ is

$$\bar{X} \sim N\left(3000, \frac{500^2}{44}\right) = N(3000, 5681.818).$$

- Now we can calculate the probability of obtaining a sample with a mean as large as 3275.955 using standardization.

Example: Baby Birth Weight

- P-value:

$$
\begin{aligned}
P(\bar{X}) > 3275.955) &= P\left(\frac{\bar{X} - 3000}{75.378} > \frac{3275.955 - 3000}{75.378}\right) \\
&= P(Z > 3.66) = 0.00015
\end{aligned}
$$

Example: Baby Birth Weight

- The p-value is very low: the probability of the data is very low if we assume the null hypothesis is true.
- Suppose that the significance level is $\alpha = 0.01$.
- In this case, we conclude that

  "there is significant evidence against the null hypothesis at the 0.01 level."

- Another way of saying this is that

  "we reject the null hypothesis at the 0.01 level."

Example: Baby Birth Weight

- The p-value is very large: the probability of the data is very high if we assume the null hypothesis is true.
- In this case, we conclude that

  "we cannot reject the null hypothesis at the 0.01 level."
- It does not mean the null hypothesis is true.