ISEN 310 Notes

Christopher Abib

December 2022

Descriptive Statistics

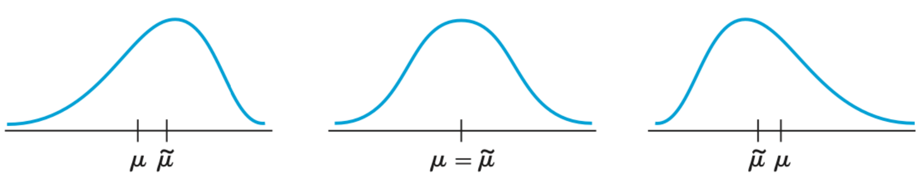
Measures of Location

The *Mean* represents the calculated center of a set of values. The *Sample Mean* ($\bar{x}$) is the point estimate of the *Population Mean* ($/mu$).

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

The *Median* represent the middling value of a set of numbers, after it has been ordered from least to greatest. The *Sample Median* ($\tilde{x}$) is the point estimate of the *Population Median* ($\tilde{\mu}$).

$$\tilde{x} = \begin{cases} n \text{ is even} & (\frac{n+1}{2})^{th}\text{ordered value} \\ n \text{ is odd} & \text{average of } (\frac{n}{2})^{th} \text{ and } (\frac{n}{2}+1)^{th}\text{ordered value} \end{cases}$$



(a) Negative skew (b) Symmetric (c) Positive skew

Measures of Variability

The *Sample Variance* (s^2) is the point estimate of the *Population Variance* (σ^2). For this estimate, the x_i's tend to be closer to the sample mean ($\bar{x}$) than to the population mean (μ). To compensate for this, the *Sample Variance* is taken with 1 degree of freedom, having a denominator of $n-1$, where n is the sample size.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

In the *Population Variance* is taken with N as the population size.

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

The *Sample Standard Deviation* (s) is the point estimate of the *Population Standard Deviation* (σ).

$$s = \sqrt{s^2}$$

$$\sigma = \sqrt{\sigma^2}$$

```
var(X)
sd(X)
```

A measure of spread that is resistant to outliers is the *Fourth Spread*, where Q_i represents each quartile. Typically, any observation farther than $1.5f_s$ from the closest quartile is an outlier. An outlier is extreme if it is more than $3f_s$ from the nearest quartile, and mild otherwise.

$$f_s = Q_3 - Q_1$$

```
quantile(X, probs=c(0, .25, .5, .75, 1))
```

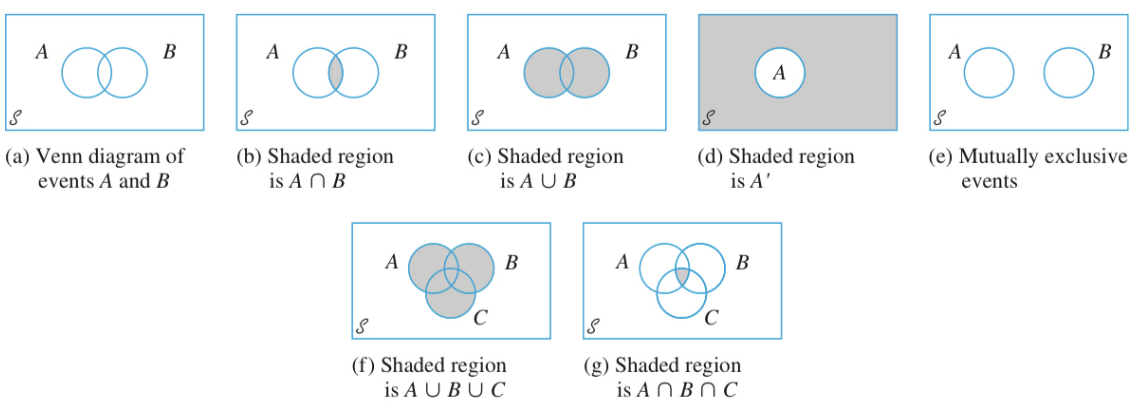
Probability

The *Sample Space* of an experiment, denoted by $\mathcal{S}$, is the set of all outcomes of that experiment.
An *Event* is any collection (subset) of outcomes contained in the sample space $\mathcal{S}$. An event is simple if it consists of exactly one outcome and compound if it consists of more than one outcome.

Some relations from set theory:

1. The *compliment* of an event A, denoted by A', is the set of all outcomes in $\mathcal{S}$ that are not contained in A.
2. The *union* of two events A and B, denoted by $A \cup B$ and read "A or B", is the event consisting of all outcomes that are in A and/or B.
3. The *intersection* of two events A and B, denoted by $A \cap B$ and read "A and B," is the event consisting of all outcomes that are in both A and B.

Let $\emptyset$ denote the *null event* (the event consisting of no outcomes whatsoever). When $A \cap B = \emptyset$, A and B are said to be *mutually exclusive* or *disjoint* events.

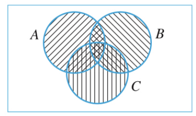


(a) Venn diagram of events A and B
(b) Shaded region is $A \cap B$
(c) Shaded region is $A \cup B$
(d) Shaded region is A'
(e) Mutually exclusive events
(f) Shaded region is $A \cup B \cup C$
(g) Shaded region is $A \cap B \cap C$

Basic probability principles:

1. *Axiom 1*: For any event A, $P(A) \geq 0$
2. *Axiom 2*: $P(\mathcal{S}) = 1$
3. *Axiom 3*: If $A_1, A_2, A_3, \cdots$ are disjoint events, then $P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$
4. $P(\emptyset) = 0$
5. For any event A, $P(A) + P(A') = 1$

6. For any event A, $P(A) \leq 1$
7. *Addition Rule*: For any two events A and B,
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
8. *Addition Rule*: For any two events A, B, and C,
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$



Conditional Probability

For any two events A and B with $P(B) > 0$, the *conditional probability* of A given that B has occurred is defined by
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
and
$$P(A \cap B) = P(A|B) \cdot P(B)$$

0.0.1 Bayes' Theorem

Law of Total Probability: Let $A_1, ..., A_k$ be mutually exclusive and exhaustive events. Then for any other event B,
$$P(B) = \sum_{i=1}^{k} P(B|A_i) \cdot P(A_i)$$
Bayes' Theorem: Let $A_1, ..., A_k$ be mutually exclusive and exhaustive events with prior probabilities $P(A_i)$ where $(i = 1, ..., k$. Then for any other event B for which $P(B) > 0$, the posterior probability of A_j given that B has occurred it
$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j) \cdot P(A_j)}{\sum_{i=1}^{k} P(B|A_i) \cdot P(A_i)} \quad j = 1, ..., k$$

Independence

Two events A and B are *independent* if $P(A|B) = P(A)$ and are *dependent* otherwise.

Multiplication Rule: A and B are independent if and only if
$$P(A \cap B) = P(A) \cdot P(B)$$
Events $A_1, ..., A_n$ are *mutually independent* if for every k $(k = 2, ..., n)$ and every subset of indices $i_1, ..., i_n$,
$$P(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_n} = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_n})$$

Counting Techniques

Letting N denote the number of outcomes in a sample space, where all outcomes are equally likely, and $N(A)$ represent the number of outcomes in an event A,
$$P(A) = \frac{N(A)}{N}$$
Product Rule: If the first element or object of an ordered pair can be selected in n_1 ways, and for each of these n_1 ways the second element of the pair can be selected in n_2 ways, then the number of pairs is $n_1 \cdot n_2$.

Combinatorics

Combinatorics is an branch of math studying the enumeration, combination, and permutation of sets of elements.

Combinations

aka: the *binomial coefficient*, choice number, n *choose* k
The number of ways of choosing k **unordered** outcomes from n possibilities.

$$_n\mathrm{C}_k = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

This code returns the binomial coefficient:

```
choose(n, k)
```

This code returns a 'k x *length(X)*' matrix where the columns are each combination:

```
combn(X, k)
```

Permutations

aka: arrangement number, order, n *pick* k

$$_n\mathrm{P}_k = \frac{n!}{(n-k)!}$$

```
choose(n, k) * factorial(k)
```

Random Variables

For a given sample space $\mathcal{S}$ of some experiment, a *random variable (rv)* is any rule that associates a number with each outcome in $\mathcal{S}$. In mathematical language, a random variable is a function whose domain is the sample space and whose range is the set of real numbers. Random variables are customarily denoted by uppercase letters; lower case letters are used to represent a particular value of the corresponding random variable. The notation $X(\omega) = x$ means that x is the value associated with the outcome ω by the rv X.

Discrete Random Variable: an rv whose possible values either constitute a finite set or else can be listed in a countably (integer) infinite sequence
Continuous Random Variable: an rv in which both of the following apply

1. Its set of possible values consists either of all numbers in a single interval on the number line (possibly infinite in extent)
2. No possible value of the variable has positive probability, that is, $P(X = c) = 0$ for any possible value c

Expected Value

$$\mu_X = \mathrm{E}[X]$$

Variance

Variance

$$\sigma_X^2 = \mathrm{Var}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2$$

Standard Deviation

$$\sigma_X = \sqrt{\mathrm{Var}[X]}$$

Variance of a Linear Function

$$\mathrm{Var}[aX+b] = \sigma_{aX+b}^2 = a^2\sigma_X^2$$

Discrete Random Variables

$$F(X) = P(X \leq x)$$

Discrete Distributions

Bernoulli Distribution

Bernoulli Random Variable: any random variable whose only possible values are 0 and 1.

Binomial Distribution

The *Binomial Random Variable* X associated with a binomial experiment consisting of n trials is defined as: $X =$ the numbers of successes among n trials where there is a p chance of each success.

$$X \sim \mathrm{Bin}(n, p)$$

$$\mathrm{E}[X] = np$$

$$\mathrm{Var}[X] = np(1-p) = npq$$

$$b(x; n, p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, 2, ..., n \\ 0 & otherwise \end{cases}$$

$$B(x; n, p) = \sum_{y=0}^{x} b(y; n, p)$$

R code to compute the distribution, where size=n and prob=p.

```
dbinom(x, size, prob) # pdf
pbinom(q, size, prob) # cdf
qbinom(p, size, prob) # quantile
rbinom(n, size, prob) # random numbers
```

Poisson Distribution

$$X \sim \mathrm{Poisson}(\lambda)$$

$$\mathrm{E}[X] = \lambda$$

$$\mathrm{Var}[X] = \lambda$$

$$f(x;\lambda) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \qquad x = 1, 2, 3, ...$$

$$F(x;\lambda) = \sum_{y=0}^{x} f(y;\lambda)$$

R code to compute the distribution, where lambda=λ

```
dpois(x, lambda) # pdf
ppois(q, lambda) # cdf
qpois(p, lambda) # quantile
rpois(n, lambda) # random numbers
```

Continuous Random Variables

Continuous Distributions

Normal Distribution

$$X \sim \mathrm{N}(\mu, \sigma^2)$$

$$\mathrm{E}[X] = \mu$$

$$\mathrm{Var}[X] = \sigma^2$$

$$f(x;\lambda) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

$$F(x;\lambda) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \qquad -\infty < x < \infty$$

R code to compute the distribution, where mean=μ and sd=σ.

```
dnorm(x, mean=0, sd=1) # pdf
pnorm(q, mean=0, sd=1) # cdf
qnorm(p, mean=0, sd=1) # quantile
rnorm(n, mean=0, sd=1) # random numbers
```

Standard Normal Distribution

Special case of the Normal Distribution

$$Z \sim \mathrm{N}(0, 1)$$

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

$$\mathrm{E}[Z] = 0$$

$$\mathrm{Var}[Z] = 1$$

$$\varphi(z;\lambda) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad -\infty < z < \infty$$

$$\Phi(z;\lambda) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{z}{\sqrt{2}}\right)\right] \qquad -\infty < z < \infty$$

Exponential Distribution

$$X \sim \mathrm{Exp}(\lambda)$$

Suppose that the number of events occurring in any time interval of length t has a Poisson distribution $X \sim Poisson(\lambda = \alpha t)$ (where α, the rate of the event process, is the expected number of events occurring in 1 unit of time) and that numbers of occurences in nonoverlapping intercals are indfependent of one another. Then the distribution of elapsed time between the occurrence of two successive events is $Y \sim Exp(\lambda = \alpha)$.

$$\mathrm{E}[X] = \frac{1}{\lambda}$$

$$\mathrm{Var}[X] = \frac{1}{\lambda^2}$$

$$f(x;\lambda) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \geq 0 \end{cases}$$

$$F(x;\lambda) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

R code to compute the distribution, where rate=λ.

```
dexp(x, rate=1) # pdf
pexp(q, rate=1) # cdf
qexp(p, rate=1) # quantile
rexp(n, rate=1) # random numbers
```

Gamma Distribution

$$X \sim \mathrm{Gamma}(\alpha, \beta)$$

The Exponential Distribution is actually a special case of the Gamma Distribution where $\mathrm{Gamma}(1,\lambda) = \mathrm{Exp}(\lambda)$.

$$\mathrm{E}[X] = \alpha\beta$$

$$\mathrm{Var}[X] = \alpha\beta^2$$

R code to compute the distribution, where shape=α and rate=$\frac{1}{\beta}$.

```r
dgamma(x, shape, rate=1) # pdf
pgamma(q, shape, rate=1) # cdf
qgamma(p, shape, rate=1) # quantile
rgamma(n, shape, rate=1) # random numbers
```

Joint Probability Distributions and Random Samples

The covariance between two rv's X and Y

$$\mathrm{Cov}[X,Y] = \mathrm{E}[X \cdot Y] - \mu_X \cdot \mu_Y$$

The correlation coefficient of X and Y denoted by $\mathrm{Corr}[X,Y]$ or $\rho_{X,Y}$

$$\mathrm{Corr}[X,Y] = \frac{\mathrm{Cov}[X,Y]}{\sigma_X \cdot \sigma_Y}$$

Point Estimation

A *point estimate* of a parameter θ is a single number that can be regarded as a sensible value for θ. It is obtained by selecting a suitible statistic and computing its value from a given sample data. The selected statistic is called the *point estimator* of θ.