**Chapter 6 Descriptive Statistics**

Let x1,…,xn denote a random sample of size n. The sample mean is μ = (x1+...+xn)/n = ∑xi /n.

Let x1,…,xN denote the observations of a population. The populationmean is μ = (x1+...+xN)/N = ∑xi /N.

Variance of a random sample: s2 = [ ∑ x**i**2 - ( ∑ x**i**)2/n] / (n-1).

Variance of population: σ2 = ∑ (x**i** - µ)2/ n.

Find Mean or Std Deviation with R: x = c(x1, x2, x3, x4, … xN)

mean(x) or sd(x)

Interquartile Range - IQR = Q3 - Q1.

Scatterplot Using R x=c(x1,x2,x3,…,xN)

y=c(y1,y2,y3,…,yN)

plot(x,y,xlab=”x”,ylab=”y”)

Describe the overall pattern of a scatterplot by the form, direction, and strength. Look for positive association (🡭) or negative association (🡮). Look for a linear pattern. How strong (striking) is the relationship? Are there clusters? Are there outliers?

**Chapter 5 Joint Probability Distributions**

If X and Y are independent random variables, then Cov(X,Y) = 0.

The ***conditional expectation*** of g(X) given Y = y is defined to be

E[g(X) | Y = y] = Σ g(x) f(x|y).

Let X1,…,Xn represent n random variables and let a1,…,an denote n constants. The random variable Y = a1X1 + … + anXn is a **linear combination** of the Xi’s.

Let X1,…,Xn denote independent rv’s where Xi has mean μi and variance σi2 for

i = 1,..,n. Then 1. μY = E(Y) = a1E(X1) + … + anE(Xn) = a1μ1 + … + anμn

2. σY2 = Var(Y) = a12 σ12 +…+ an2 σn2.

1) Let X and Y be jointly discrete with joint density given by the table below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Y=1 | Y=2 | Y=3 | Y=4 | fx(x) |
| X=0 | 0.05 | 0.10 | 0.15 | 0.05 | 0.30 |
| X=1 | 0.05 | 0.05 | 0.10 | 0.05 | 0.25 |
| X=2 | 0.20 | 0.05 | 0.05 | 0.10 | 0.40 |
| fy(y) | 0.30 | 0.20 | 0.30 | 0.20 | 1 |

1. Marginal densities fy(y) and fx(x) are found by adding rows or columns
2. COV(X,Y) is found with E(XY) - E(XY)- μxμy

μx = Σxi fx(x) μy = ΣYjfy(y) E(XY) = XiYjP(X = i, Y = j)

1. Probability within a range: P(X < 2, Y >3) = summation of the probability at each coordinate where X < 2 and Y > 3
2. Expected value of Y given X = 0 -> E(Y=y | X=0) = (1(.05) + 2(0.1) + 3(.15) + 4(.05))(100/35)

We multiply by 100/35 to adjust the total probability of the sample size.

1. Are X and Y independent. Only if f(X,Y) = fy(y)fx(x) at every possible occurrence.

2) Let X and Y be jointly continuous with joint density f(x,y) = (5/81)x2y, 0 < x < y < 3

1. marginal density of y -> fy(y) = integral from 0 to y of f(x,y)dx. Define end limit for fy(y) which was 0 < y < 3 in this case.
2. E(XY) = integral from 0 to 3 of integral from 0 to y of xyf(x,y)dxdy
3. conditional density of X given Y = y. h(x|y) = f(x,y)/ fy(y) and make sure to define the interval. In this case it was 0 < x < y.
4. Are X and Y independent. Only if f(X,Y) = fy(y)fx(x).
5. To combine standard deviations, square both, add, then take square root.

**Chapter 5 Joint Probability Distributions**

Joint Probability Mass Function - Let X and Y be discrete random variables. The ***joint probability distribution*** for X and Y is given by f(x,y) = P(X = x, Y = y). The function f(x,y) can be called the ***joint probability mass function*** for X and Y. f(x,y) ≥ 0 Σ f(x,y) = 1

Let X and Y be jointly discrete random variables with probability function f(x,y). The ***marginal probability mass functions*** of X and Y, respectively, are given by fX(x) = Σ f(x,y) in 1 x col and fY(y) = Σ f(x,y) in 1 y col.

The ***conditional probability mass function*** of X given Y is given by

f(x | y) = P(X = x | Y = y) = P(X = x, Y = y)/P(Y = y) = f(x,y)/fY(y) , fY(y) > 0.

X and Y are ***independent*** random variables iff f(x,y) = fX(x) fY(y).

Let g(X,Y) be a function of the random variables X and Y. Expected Value E[g(X,Y)] = ΣΣg(x,y)f(x,y).

Properties of Expected Values : 1. Let c be a constant. Then E(c)= c.

2. E[cg(X,Y)] = c E[g(X,Y)]

***Covariance*** is a measure of dependence. If X and Y are independent, then the covariance between X and Y is 0. Let μ1 = E(X) and μ2 = E(Y). The ***covariance*** of X and Y is COV(X,Y) = E[(X -μ1)(Y - μ2)] = E(XY)- μ1μ2

Population linear correlation coefficient is ρ = Cov(X,Y) / (σ1 σ2)

Properties of ρ 1. -1 ≤ ρ ≤ 1

2. ρ = 0 ⇒ no correlation between X and Y

3. ρ ≈ 1 ⇒ strong, positive correlation btwn X and Y

4. ρ ≈ -1⇒ strong, negative correlation btwn X and Y

**Chapter 4 Continuous RV**

Normal Approx. to Binomial : Assume that X is a Binomial(n,p) rv. If np and n(1 - p) are both at least 10, then X is approximately N(np, np(1 - p)).

Continuity Correction : either add or subtract 0.5 of a unit from each discrete x-value.

Exponential Distribution : An Exponential(λ) rv X has pdf

f(x) = λe-λx , x > 0 (λ > 0) and E(X) = 1/λ and V(X) = 1/λ2.

**Chapter 4 Continuous RV**

The mathematical model for a continuous rv X is a **probability density function** (pdf) whose graph is a **probability density curve**.

Let f(x) denote the pdf of a continuous rv X.

NOTE: (1) f(x) ≥ 0

(2) f(x) dx = 1 b

(3) P(a < X < b) = ∫a f(x) dx

(4) P(X = a) = 0

(5) P(a ≤ X < b) = P(a ≤ X ≤ b) = P(a < X ≤ b) = P(a < X < b)

X is a **Uniform(a,b)** rv along interval a to b iff f(x) = 1/(b - a), a < x < b.

For the Uniform(a,b), E(X) = (a + b)/2 and V(X) = (b – a)2/12.

The cdf of X is F(x) = P(X ≤ x). In the continuous case, this is computed by



Mean = μ = E(X) = ∫ x f(x) dx.

Variance = V(X) = σ2 = σ2X = E[(X-μ)2] = ∫(x-μ)2 f(x)dx.

The **standard deviation** of X is given by σX, which is the square root of σ2X.

**Standard Normal RV :** Z has a N(0, 1) (or standard normal) distribution when the pdf of Z is the following:

**f(z) = , -∞ < z < ∞**

****

Percentile : z-score gives percentile. z = (x - μ) / σ

To find P(Z ≤ x) with R: pnorm(x)

To find P(X < x) for a N(a,b): pnorm(x,a,b)

To find xth percentile of N(0,1) : qnorm(0.x)

To find the xth percentile of a N(a,b): qnorm(0.x,a,b)

**Chapter 3 Discrete RV**

When f(x) = P(X = x) depends on a quantity which can take on various values, we call this quantity a **parameter**.

The **cumulative distribution function (cdf)** of a rv X is F(x) = P(X ≤ x).

**Properties of F**

1. F(- ∞) = 0 and F(∞) = 1

2. For a < b, F(a) ≤ F(b)

3. F is continuous from the right.

Because F(x) is a probability, 0 ≤ F(x) ≤ 1.

Let X be a discrete rv. The **expected value** (or **expectation** or **mean**) of X is given by μ = μX = E(X) = Σ x P(X = x). Or let h be a function of X, where X is a discrete rv. Then E[h(X)] = Σ h(x) P(X=x).

E(aX + b) = aE(X) + b, where a and b are constants.

Let X be a discrete rv. The **variance** of X is given by

V(X) = σ2 = σ2X = E[(X-μ)2] = Σ(x-μ)2 f(x) = E(X2) - (E(X))2 = Σ x2 f(x) - μ2

V(aX + b) = a2 V(X)

**Standard deviation** of X given by σ or σX, which is the square root of σ2.

Discrete Uniform DistributionAssume that the random variable X assumes values x1, x2, …, xn with equal probability. Then f(xi) = P(X = xi) = 1/n

and E(X) = (n + 1)/2 and V(X) = (n2 – 1)/12.

Binomial Random Variable. X is a Binomial(n,p) rv with pmf:

P(X = x) = n! px (1 - p)n-x , x = 0,1,…,n.

x!(n - x)!

dbinom(x,n,p) computes P(X = x) for a Bin(n,p) rv.

pbinom(x,n,p) computes P(X ≤ x) for a Bin(n,p) rv

x -> # of successes, n -> # of attempts, p -> probability of success

Binomial(n,p) Mean = E(X) = np

Binomial(n,p) Variance = V(X) = np(1 - p).

**Chapter 2 Basic Probability of Set Relations**

The **sample space**, S, is the set of all possible outcomes of an experiment.

An **event** is a subset of S or a collection of outcomes of an experiment. Each outcome is a **simple event**.

A∪B = A or B

A∩B = A and B

A**′** = complement of A = not A

A and B are **mutually exclusive** or **disjoint** if and only if A∩B = ∅.

A sample space is [**discrete**](http://edugen.wileyplus.com/edugen/courses/crs7600/ebook/c02/ebook/c02/montgomery9781118539712c02xlinks.xform?id=c02-term-2008) iff it consists of a finite or countable infinite set of outcomes.

A sample space is [**continuous**](http://edugen.wileyplus.com/edugen/courses/crs7600/ebook/c02/ebook/c02/montgomery9781118539712c02xlinks.xform?id=c02-term-2009) iff it contains an interval (either finite or infinite) of real numbers.

A **U** B = B **U** A and A ∩ B = B **∩** A

(A **U** B) **∩** C = (A **∩** C) **U** (B **∩** C)

(A **∩** B) **U** C = (A **U** C) **∩** (B **U** C)

(A **U** B) **U** C = A **U** (B **U** C)

(A **∩** B) **∩** C = A **∩** (B **∩** C)

(A **U** B)**′** = A**′** **∩** B**′** and (A **∩** B)**′** = A**′** **U** B**′**

(A**′**)**′** = A

P(A) = probability of an event A

*Axiom 1* P(A) ≥ 0 for any event A.

*Axiom 2* P(S) = 1.

*Axiom 3* Let A1,…,An be n mutually exclusive events. Then P(A1∪...∪An) = P(A1) +…+ P(An).

*Long term relative frequency -* If you toss a coin a large number of times, say N, you would expect number of heads out of N ≈ 0.5.

Law of Complementation: For an event A, P(A) = 1 - P(A**′**)

Disjoint : If A and B are disjoint (i.e., A∩B = ∅), then P(A∩B) = 0

Addition Rule: For any events A and B, P(A∪B) = P(A) + P(B) - P(A∩B).

*Conditional Probability* of A given B has occurred is given by

P(A|B) = P(A∩B)/P(B) , P(B) > 0

Multiplication Rule - P(A|B) = P(A∩B)/P(B) ⇒ P(A∩B) = P(A|B)P(B)

P(A∩B) = P(A|B)P(B), P(B) > 0

or P(A∩B) = P(B|A)P(A), P(A) > 0

A and B are **independent** iff one of the following equivalent statements hold: (i) P(A∩B) = P(A)P(B);

(ii) P(A|B) = P(A), P(B) > 0;

(iii) P(B|A) = P(B), P(A) > 0.

Theorem of Total Probabilities : Assume E1, E2, … , Ek are mutually exclusive events such that S = UEi . For an event A, P(A) = P(A|E1)·P(E1) + … + P(A|Ek)·P(Ek)

Baye’s Theorem : we observe conditional probabilities through prior information. P(E1 |A) = P(A| E1) P(E1) .

P(A|E1)·P(E1) + … + P(A|Ek)·P(Ek)

A **random variable** (rv) is a function from the sample space into the real numbers. In other words, a random variable assigns a number to each outcome of an experiment.

**Chapter 1 Intro Definitions**

Descriptive Statistics - part of statistics dealing with organizing and summarizing data

Inferential Statistics - part of statistics dealing with drawing conclusions about the population on the basis of a sample

Probability - knowing the population, probability calculations can be made concerning a sample.

Statistics - given the sample, we infer about the population.

A **retrospective study** using historical data. Data collected in the past for other purposes.

An **observational study**. Data, presently collected, by a passive observer.

A **designed experiment.** Data collected in response to process input changes. Treatments are deliberately imposed to observe a response.

A **simple random sample (SRS)** of size *n* is a sample that has been selected from a population in such a way that each possible sample of size *n* has an equally likely chance of being selected.

The value of a ***categorical* variable** depends on which of several categories an individual falls.

The value of a ***quantitative* variable** is obtained by a measurement or a count for an individual.

