=sample mean

**Linear Transformation**

Constants a and b, if

…then Ex temperature in F to C, compute average in F, then convert that average to C

=sample variance=

**Elements of Probability**

Sample space S: set of all possible outcomes

Event , Any subset of sample set

Union: , E || F

Intersect: , E && F, EF

Null event: has no outcome

Compliment: *,* , all outcomes in but not in

 and are mutually exclusive,

**Containment** 𝐸⊂𝐹 or 𝐹⊃𝐸

all out comes in 𝐸 are also in 𝐹

Example: 𝐸={(𝐻,𝐻)} ⊂ 𝐹={(𝐻,𝐻),(𝐻,𝑇), (𝑇,𝐻)}

Occurrence of 𝐸 implies that of 𝐹

Example: getting two heads implies getting head at all

Equality

𝐸=𝐹, if 𝐸⊂𝐹 and 𝐸⊃𝐹

**Commutative law** 𝐸∪𝐹=𝐹∪𝐸, 𝐸𝐹=𝐹𝐸

**Associative law** (𝐸∪𝐹)∪𝐺=𝐸∪(𝐹∪𝐺), (𝐸𝐹)𝐺=𝐸(𝐹𝐺)

**Distributive law** (𝐸∪𝐹)=𝐸𝐺∪𝐹𝐺, 𝐸𝐹∪𝐺=(𝐸∪𝐺)(𝐹∪𝐺)

**DeMorgan’s law** (𝐸∪𝐹)^𝑐=𝐸^𝑐 𝐹^𝑐, (𝐸𝐹)^𝑐=𝐸^𝑐 𝐹^𝑐

**Axioms of Probability**

Probability of event 𝐸=Proportion of times 𝐸 occurs (out come is in 𝐸) in repeated experiments

Axiomatic definition: For each event 𝐸 of an experiment with sample space 𝑆, assign a number (𝐸). If (𝐸) satisfies the following three axioms

0≤(𝐸)≤1, 𝑃(𝑆)=1

For any sequence of mutually exclusive events 𝐸\_1,\_2,…,

 We call (𝐸) the probability of event 𝐸

Example:

Proof

 and are mutually exclusive,

Example:

Proof: Use Venn diagram

*Odds* of event

, odds, it is equally likely

, odds = 3, it is 3 times as likely to occur

**Conditional Probability**

 ,

**Law of Total Probability**

Proof:

*,*

 and are mutually exclusive

**Bayes’ Formula**

**Independent Events**

, if

**Discrete r.v.**

Pmf =

Cdf =

**Continuous r.v.**

Pdf =

Cdf =

**Joint r.v.**

*joint CDF*,

Discrete RVs, *joint PMF*,

*Marginal CDF* , i.e. can take any value

*Marginal PMF*

Continuous joint PDF:

Joint CDF:

*Marginal PDF*

**Independent r.v.** , if for any two sets of real numbers and ,

If :

, for discrete RVs

, for continuous RVs

**Conditional Distribution**

Disc:

Continuous:

**Expectation** discrete RV:

Continuous RV:

If independence,

To solve optimization problems:

-choose quantity to represent r.v.

MSE=Mean Squared Error,

Let that minimizes MSE,

**Law of the Lazy Statistician**

 unknown, but known

 except for linear case

**Variance**

-always positive, 0 iff is constant

**Covariance**

 ,

If , , but converse may be false

**Variance and Covariance**

Variance of sum of RVs

If ,

For , if , for all ,

Mutually independent

*Correlation*:

, correlation does not mean independence

**Markov’s Inequality** is a nonnegative random variable, then for any ,

Proof:

**Chebyshev’s Inequality** is a random variable with mean and variance , then

Proof: Consider random variable , which is nonnegative. Apply Markov’s inequality with

**Weak Law of Large Numbers**

Let be a sequence of independent and identically distributed (i.i.d.) random variables, each having mean . Then for any ,

Average of i.i.d. RV’s converges to their mean

Proof: (assuming exists)

Apply Chebyshev’s inequality

, as

**Bernoulli**

**Binomial** sum of iid Bernoulli r.v.

, for

Ex. Acceptance Sampling

Given an Operating Characteristic curve, want to know probability of acceptance/rejection

Split into 4 regions: accept acceptable, reject nonacceptable (hit), reject acceptable (supplier’s risk, false alarm/positive, type I error), accept nonacceptable (inspector’s risk, miss, type II error)

As number of products inspected go up, curve gets steeper, ideally a step function

**Poisson**

Number of event occurrences during a time period

Example: arrival of customers

We observe that on average customers come to the store in 1 hour

Divide 1 hour into intervals (e.g. 3600 seconds)

On average, out of intervals have customer arrive

When is large, no interval has more than 1 arrival in it

Each interval has arrival with probability

Number of arrivals in one hour: binomial distribution

,

Pois replaces bin when large, small

,

**Uniform** If we want the range to be

consider

CDF: ,

, if

PDF:

Mean:

Variance:

**Exponential** has PDF , for

CDF

For a non-negative random variable , with tail distribution

**Normal** PDF

If , , then

, then , standard normal dist

CDF:

:

The -th percentile of

Sum of independent normal RVs is still normal

, ,

What about ?

**Samples** sample mean

Sample variance

For an arbitrary distribution with mean and var

For the sample mean ,

Approximate distribution when is large

(Central Limit Theorem)

For the sample variance ()

For a normal distribution , Exact distribution of and

**Central Limit Theorem** Let be a sequence of i.i.d. random variables with mean and variance

then the distribution of converges to standard normal as

Generally works as long as

Questions: approximate distribution of the sum?

Question: approximate distribution of the sample mean?

**Ex** The College of Engineering decided to admit 450 students, 0.3 probability that an admitted student will actually enroll. What is the probability that the incoming class has more than 150 students?

Solution:,

 is approximately normal with mean and standard deviation

Approximate a discrete r.v. (e.g. ) with a continuous r.v. (e.g. )

Question:

*Continuity correction*

**Continuity Correction** When using a continuous function to approximate a discrete one, ex. Normal approx.. Bin.

If P(X=n) use P(n – 0.5 < X < n + 0.5)

If P(X>n) use P(X > n + 0.5)

If P(X≤n) use P(X < n + 0.5)

If P (X<n) use P(X < n – 0.5)

If P(X ≥ n) use P(X > n – 0.5)

**Sample Variance**

, so unbiased

**Sampling from a Normal Population**

**Chi-square distribution** with degrees of freedom, sum of indep chi-squar is chi-square

,

Rearrange the terms

Decomposition of the sum of squares

Divide by

=

, are independent standard normal

 is chi-square with degrees of freedom

If sampling from normal population, then , sample mean and variance are indep r.v.

CPU time for a type of jobs is normally distributed with mean seconds and standard deviation seconds

 such jobs are tested, what is the probability that the sample variance exceeds

Solution

**-distribution** If with deg of freedom Recall

**Parameter Estimation** a population has certain distribution with unknown parameter

Estimate based on

 a function of sample, is a r.v.

Is also a point estimator

**Moment Generating Function**

 moment is derivative

**Method of Moments**

Moments can be expressed as functions of parameters

If unknown parameters,

**Maximum Likelihood Estimators**

Given parameter , for each observation , we can calculate the joint density function

 are independent

Given observation , for each , define the *likelihood*

Maximum likelihood Estimator MLE maximizes

For MLE use the pdf of the distribution, using to replace as needed. Multiply together all samples in this new pdf.

It’s easier to maximize

MLE of exponential dist is the sample mean

Example: Bernoulli distribution

Normal:

 ,

Uniform on [0,],

Poisson:

**Point Estimators**

Biased/unbiased and precise/imprecise

*Bias* of estimator for parameter

**Confidence Interval**

2-sided, 1-sided upper/lower

Ex. normal 2-sided 95% confidence

If is given, and we have observed sample mean

 is a  *confidence interval estimate* of

Once observed, is a constant, and the interval is fixed. is not a probability

**Two-sided Confidence Interval**  for normal mean with known variance:

One sided upper:

One sided lower:

**Estimating Difference in Means**

Normalize

 is a

C.I. for the difference in the means of two normal population when the variances are known

When unknown variances, assume to be equal

Use Pooled Sample Variance

 is a C.I. for the difference in the means of two normal population when the variances are unknown but equal

**Approximate Confidence Interval**

If population not normal but relatively large sample size, use results for normal population mean with known variance to obtain approximate c.i.

Ex. Monte Carlo Simul.

Evaluate a definite integral

Consider an uniform random variable in

Law of lazy statistician

Estimate

Simulate uniform random variables

Let and be the observed sample mean and standard deviation of the ’s

Approximate C.I.