STAT 400
Study Guide

Donald Isaac

September 11, 2020

1 Introduction

This study guide contains formulas and brief explanations for concepts from STAT 400 at the University of Maryland.

2 Probability

2.1 Sample Spaces and Events

$$\forall E_k \in S, \quad P(E_k) = \frac{N(E_k)}{N(S)}$$

$$n := |S|, \quad P(S) = \sum_{k=0}^{n} P(E_k) = 1$$

2.2 Properties of Probability

$$P(E^c) = 1 - P(E)$$

$$P(B \cap C) = P(A \cap B \cap C) + P(A^c \cap B \cap C)$$

When events A, B are independent:

$$P(A \cup B) = P(A) + P(B) - P(A \cup B)$$

$$P(A \cap B) = P(A)P(B)$$

When disjoint:

$$P(a \cup B) = P(A) + P(B)$$

$$P(A \cap B) = \phi$$

2.3 Permutations and Combinations

$${}_nP_k = \frac{n!}{(n-k)!}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

2.4 Conditional Probability & Bayes' Theorem

The probability of event A occurring, given that B has occurred, is

$$P(A|B) = \frac{P \cap B}{P(B)} = \frac{P(B|A)}{P(B)} P(A)$$

When A and B are independent, this reduces to

$$P(A|B) = P(A)$$

2.5 Independence of Events

3 Discrete Random Variables and Probability Distributions

3.1 Random Variables

Random Variable (rv): A variable measuring some characteristic of an experiment's outcomes and is usually denoted as X. Can be discrete or continuous
Bernoulli Random Variable: A discrete rv who's value can only be 0 or 1.

3.2 Probability Distributions

Probability Mass Function (pmf): The *pmf* of X specifies the probability of observing a specific outcome value x of an experiment. More formally, the *pmf* $p(x)$ is defined for all x such that

$$p(x) = P(X = x) = P(\forall \gamma \in S : X(\gamma) = x)$$

Cumulative Distribution Function (cmf): The *cmf* $F(x)$ for a discrete rv X is the probability that X will be at most x.

$$F(x) := P(X \leq x) = \sum_{y|y \leq x} p(y)$$

3.3 Expected Value, Variance, and Std. Deviation

Expected Value: The *expected value* of the discrete rv X is it's average value. If the set of all possible outcomes of X is V, and outcome $x \in V$ has a value function $h(x)$, then

$$E(X) = \mu_X = \sum_{x \in V} h(x)p(x)$$

Variance: Expresses the amount of variability for values of X.

$$V(X) = \sum_{x \in V} (h(x) - \mu_X)^2 p(x) = E\big[(X - \mu_X)\big]$$

x can be substituted for $h(x)$ when $h(x) = x$.
Standard Deviation: The square root of the variance.

$$\sigma_X = \sqrt{V(X)}$$

Using the formula for the standard deviation, the variance formula can be reduced to

$$V(X) = \sigma_X^2 = \Big[\sum_{x \in V} x^2 * p(x)\Big] - \mu_X^2 = E(X^2) - [E(X)]^2$$

When $h(x)$ is linear such that $h(x) = aX + b$, the expected value formula can be reduced to

$$E(h(x)) = a * E(X) + b$$

And variance / std. deviation can be reduced to

$$V(h(x)) = a^2 * \sigma_X^2 \rightarrow \sigma h(x) = |a| * \sigma_X$$

3.4 The Binomial Probability Distribution

Binomial Experiment: An experiment is a binomial experiment if: it consists of a fixed number of *trials* n; it results in one of two possible outcomes, denoted as S and F; each trial is independent from one another; and the probability of success $p = P(S)$ is constant across each trial.
Binomial Distribution: The approximate probability model for a sampling without replacement from a population of n Bernoulli trial outcomes. Let the outcome of a S trial with a probability p be denoted as the *binomial variable* X. Then, the *pmf* of X $b(x; n, p)$ is

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

Which is (*the number of n-length sequences consisting of x S's*) times (*the probability of such a sequence*). When X is a binomial rv for an experiment with n trials, each with a S probability p, it is denoted as $X \sim Bin(n, p)$. The *cdf* for $X \sim Bin(n, p)$ is

$$F(x; n, p) = P(X \leq x) = \sum_{y=0}^{x} b(y; n, p) \; \forall x \leq n$$

furthermore, the expected value, std. deviation, and variance of X if $X \sim Bin(n, p)$ is

$$E(X) = np, \; V(X) = np(1-p) \rightarrow \sigma_X = \sqrt{npq}$$

where $P(F) = q = 1 - p$.

3.5 Hypergeometric and Negative Binomial Distributions

A **Hypergeometric Distribution** of a discrete random variable X discribes the probability of k successes in n draws, without replacement.
The hypergeometric distribution pmf is

$$P(X = k) = \frac{\binom{a}{k}\binom{n-a}{r-k}}{\binom{n}{r}}$$

3.6 Poisson

A discrete random variable X is said to follow a Poisson distribution with parameter μ, if it has probability distribution

$$P(X = x) = p(x; \mu) = \frac{e^{-\mu} * \mu^x}{x!}$$

Note that $E(X) = V(X) = \mu$
$\mu = \lambda =$ average number of events. λ is the poisson constant. When given a time rate r for x events to happen, then $\lambda = rt$ and

$$P(X = x) = p(x; r, t) = e^{-rt}\frac{(rt)^x}{x!}$$

4 Continuous Random Variables and Probability Distributions

4.1 Continuous Random Variables

The **probability distribution** or **probability density function** for a continuous rv X is a function $f(x)$ such that

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

4.2 Cumulative Distribution Functions and Expected Value

4.3 Normal Random Variables

A continuous rv X with an expected value μ and standard deviation σ is said to be *Normally Distributed* when it's pmf matches the following formula

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

The special case when $\mu = 0$ and $\sigma = 1$ is called the *Standard Normal Distribution* and has a pmf of

$$\Phi(z) = f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

The Z score is the number of standard deviations an rv X is from the mean.

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

The cdf of a non-standard distribution can be found by converting X to Z

$$P(a \leq X \leq b) = P(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma})$$

$$= \Phi(\frac{a-\mu}{\sigma}) - \Phi(\frac{b-\mu}{\sigma})$$

4.4 Exponential and Gamma Distributions

A continuous rv X has an *Exponential Distribution* if the pdf of X is

$$\forall x \in \mathbb{R}^+,\ f(x;\lambda) = \lambda e^{-\lambda x}$$

and has the following expected value, variance, and standard deviation formulas

$$\mu = \sigma = \frac{1}{\lambda},\ \ V(X) = \sigma^2 = \frac{1}{\lambda^2}$$

The cdf for the exponential distribution pdf is

$$\forall x \in \mathbb{R}^+,\ F(x;\lambda) = 1 - e^{-\lambda x}$$

5 Joint Probability

5.1 Join Probability Distributions and Random Samples

5.2 Expected Values, Covariance, an Correlation

Let X and Y be jointly distributed rvs with a pmf $p(x, y) = P(X = x^Y = y)$ when they are discrete, or pdf $f(x, y)$ when continuous. Then the expected value is

$$E[h(x,y)] = \sum_x \sum_y h(x,y)p(x,y)$$

if X, Y are discrete or

$$E[h(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)dxdy$$

if they are continuous.
Covariance measures how strongly correlated X and Y are to each other. The formula for Covariance is

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

When X and Y are discrete, this turns into

$$\sum_x \sum_y (x-\mu_X)(y-\mu_Y)p(x,y)$$

or

$$E[h(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)f(x,y)dxdy$$

if they are continuous. This can be further reduced into the form

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

The *Correlation Coefficient* of X, Y, denoted as $Corr(X, Y)$ or $\rho_{X,Y}$ is

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X * \sigma_Y}$$

5.3 TODO

5.4 Sampling Distributions and the Central Limit Theorem

Given a population with an expected value μ_X and standard deviation σ_X, the *Central Limit Theorem* states that, for a sampling distribution with sample size n:

- as n approaches infinity, the sampling distribution approaches a normal distribution;
- The expected value is $E(\bar{X}) = \mu_{\bar{X}} = \mu_X$;
- The standard deviation is $\sigma_{\bar{X}} = \dfrac{\sigma_X}{\sqrt{n}}$

There are a few requirements/restrictions for using the Central Limit Theorem. Let $X_1, X_2, X_3, ..., X_i$ be the elements in the sample of size n. Then,

- Can only be used for $n \geq 30$;
- Each X_i must be independent from each other;
- Each X_i must have the same pdf

6 Point Estimation

6.1 Point Estimators

A point estimator $\hat{\theta}$ is said to be an unbiased estimator of θ if $E(\hat{\theta}) = \theta$
The point estimator for σ is

$$\hat{\sigma}^2 = S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} = \frac{1}{n-1}\Big[\sum X_i^2 - \frac{(\sum X_i)^2}{n}\Big]$$

The standard error of a point estimator is

$$\hat{\sigma} = \frac{\sigma}{\sqrt{n}}$$

for the point estimator $\hat{p}$, the standard error is

$$\hat{\sigma} = \sqrt{\frac{pq}{n}}$$

6.2 Methods of Point Estimation

Given the rvs $X_1, X_2, X_3, ..., X_n$ and $k > 0$, the kth population moment is

$$m_k(X) = E(X^k)$$

7 Chapter 7

α = probability of error, $1 - \alpha$ = confidence interval. Assuming the population has a normal distribution with a known σ,

$$P[\mu - z_{\alpha/2} < \bar{X} < \mu + z_{\alpha/2}] = 1 - \alpha$$

$$P\Big[\bar{X} - z_{\alpha/2}(\frac{\sigma}{\sqrt{n}}) < \mu < \bar{X} + z_{\alpha/2}(\frac{\sigma}{\sqrt{n}})\Big] = 1 - \alpha$$

