STAT400 Notes

Table of contents

1. Basic Terms

Parameters Quantitative features of a population

Statistics Quantitative features calculated using a sample

Deduction A way of going from the population to the sample, e.g., if you get the class average (population) and then compare it to your personal score on an exam (sample), you can deduce that you did better than the average.

Inference A way of going from the sample to the population (opposite direction from deduction). For example, if you get a 100 on an exam (sample) and you infer that you did better than the class average (population).

Inference and deduction goes in a circle. When developing a model, you estimate some parameters, then you make deductions about the population, then you infer population parameters, then you draw deductions about the population again.

Experiment Repeatable task with well-defined outcomes

Sample space The sample space for an experiment is the set of all possible outcomes of that experiment, denoted by S

Event Any subset E of the sample space S attached to an experiment will be called an event associated with the experiment. To say that an event E has happened means that the outcome of the experiment was in the set E.

Simple event An event with only one event in it

2. Probability functions

2.1. Sigma algebra

A sigma algebra is a collection of interesting events in some sample space.

A collection $\mathcal{B}$ of subsets of a sample space S is a sigma algebra if:

1. $\emptyset \in \mathcal{B}$
2. $\forall A \in \mathcal{B}, A^c \in \mathcal{B}$
3. If $\{A_i : i \in \mathbb{N}\}$ is a countable collection such that $A_i \in \mathcal{B}$ for all i, then $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{B}$

2.2. Probability functions

Consider a sample space S with a sigma algebra $\mathcal{B}$.

A probability function is a function from events to probabilities ($\mathcal{B} \to \mathbb{R}$). It must satisfy the following axioms:
1. (finite measure) $P(S) = 1$
2. (positivity) $\forall A \in \mathcal{B}, P(A) \geq 0$
3. (countable additivity) For $A_1, A_2, A_3, ...$, the collection of pairwise disjoint subsets of S in $\mathcal{B}$, we must have

$$P\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

3. Independent events

Two events A and B are independent if any of the following are true (all are equivalent):
- $P(A \cap B) = P(A)P(B)$
- $P(A \mid B) = P(A)$
- $P(B \mid A) = P(B)$

4. Random variables

A random variable X maps outcomes in some sample space to real numbers, i.e., $X : \mathcal{S} \to \mathbb{R}$. A random variable measures a specific quantitative feature of the sample space outcome.

The **range** of X, the set of all possible values that X can take, is denoted $\mathcal{X}$.

With the new sample space, $\mathcal{X}$, you can use the order relationship in real numbers and you can add, multiply, etc.

5. Discrete distributions

5.1. Geometric distribution

6. Continuous distributions

6.1. Normal distribution

todo

6.2. Gamma distribution

The gamma distribution is based on the gamma function, which extends the factorial function to complex numbers:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} \, \mathrm{d}x$$

Some properties of Γ:
- $\Gamma(x) = x!$ if x is a non-negative integer
- $\Gamma(a+1) = a\Gamma(a)$

If you have a gamma distribution $\mathrm{Gamma}(\alpha, \lambda)$, then its probability density function is

$$f_{X(x)} = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & \text{if } x > 0 \\ 0 \text{ otherwise} \end{cases}$$

Note: there are two different parameterizations for Gamma:
- Gamma$(\alpha, \beta/\lambda)$ (with a rate parameter)
- Gamma(α, θ) (with a scale parameter)`

The α is the shape parameter. $\beta = \frac{1}{\theta}$

6.3. Exponential distribution

This is a special case of the gamma distribution: Exponential$(\lambda) = $ Gamma$(1, \lambda)$. If you sum n independent Exponential(λ) random variables, you get a Gamma(n, λ) random variable.

Application: Variables that model the amount of time you have to wait before something happens follow an exponential distribution. E.g., time between clicks of a Geiger counter.

This is the continuous analog of the geometric distribution. Only the geometric and exponential distributions are **memoryless**.

6.4. Chi-squared distribution

Also a special case of the gamma distribution: $\chi^2_k \sim$ Gamma$\left(\alpha = \frac{k}{2}, \theta = 2\right)$.

todo find out how to use it