THEORY OF STATISTICS I

EDWARD O'CALLAGHAN

Contents

1. Sigma Fields

Definition 1.1 (Sigma Field)**.** A σ-field, or σ-algebra, is some subset $\mathcal{A} \subseteq 2^X$ of the power set 2^X of set X such that;

- $\emptyset, X \in \mathcal{A}$,
- if $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$,
- for any countable collection of sets $A_1, A_2, \dots \in \mathcal{A}$ we have, $\bigcap_i A_i \in \mathcal{A}$ and $\bigcup_i A_i \in \mathcal{A}$.

Lemma 1.2. *The σ-field $\mathcal{F}$ from some set Ω, taken as $\mathcal{F} = \{\emptyset, \Omega\}$, is the* trivial *$\sigma$-field.*

Definition 1.3 (Measure)**.** A function μ defined on some σ-field $\mathcal{A}$ is said to be countably additive and nonnegative *measurable* provided;

- $\mu(\emptyset) = 0$,
- $0 \leq \mu(A) \leq \infty$ for any $A \in \mathcal{A}$,
- for a countable collection of sets $A_1, A_2, \dots \in \mathcal{A}$ with $A_i \cap A_j = \emptyset : i \neq j$, we have $\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i)$.

Definition 1.4 (Measure Space)**.** The tuple $(X, \mathcal{A})$ is called a *measurable space* for some σ-field define on some set X. That is, a space in which we may define some *measure map* $\mu : \mathcal{A} \to \mathbb{R}$.

Theorem 1.5. *Let μ be a measure on $(\Omega, \mathcal{F})$*

i.) ***Monotonicity**: If $A \subset B$ then $\mu(A) \leq \mu(B)$.*
ii.) ***Subadditivity**: If $A \subset \cup_{k=1}^{\infty} A_k$ then $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$.*
iii.) ***Continuity from below**: If $A_i \uparrow A$ (i.e., $A_1 \subset A_2 \subset \dots$ and $\cup_i A_i = A$) then $\mu(A_i) \uparrow \mu(A)$.*
iv.) ***Continuity from above**: If $A_i \downarrow A$ (i.e., $A_1 \supset A_2 \supset \dots$ and $\cap_i A_i = A$), with $\mu(A_1) \leq \infty$ then $\mu(A_i) \downarrow \mu(A)$.*

Proof.

i.) Let $B - A = B \cap A^c$ be the *difference* of the two sets. Using $\oplus$ to denote disjoint unions, $B = A \oplus (B - A)$ so
$$\mu(B) = \mu(A) \oplus \mu(B - A) \geq \mu(A).$$

ii.) Let $A'_n = A_n \cap A, B_1 = A'_1$ and for $n > 1$, $B_n = A'_n - \cup_{k=1}^{n-1}(A'_k)^c$. Since the B_n are disjoint and have union A, we have, using the definition of measure, $B_k \subset A_k$, and by *i.*) of this theorem,
$$\mu(A) = \sum_{k=1}^{\infty} \mu(B_k) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

iii.) Let $B_n = A_n - A_{n-1}$. Then the B_n are disjoint and have $\cup_{k=1}^{\infty} B_k = A$, $\cup_{k=1}^{n} B_k = A_n$ so

$$\mu(A) = \sum_{k=1}^{\infty} \mu(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(B_k) = \lim_{n\to\infty} \mu(A_n).$$

iv.) Given that $A_1 - A_n \uparrow A_1 - A$ so *iii.)* implies $\mu(A_1 = A_n) \uparrow \mu(A_1 - A)$. Now since $A_1 \supset B$, we have $\mu(A_1 - B) = \mu(A) - \mu(B)$ and it follows that $\mu(A_n) \downarrow \mu(A)$.

□

Definition 1.6 (Probability Measure)**.** A *probability measure* is some measure $\mathbb{P}$ defined on some σ-field $\mathcal{F}$ from some set Ω such that $\mathbb{P}(\Omega) = 1$. In particular, we call the set Ω the *sample space* and the triple $(\Omega, \mathcal{F}, \mathbb{P})$ the *probability* measure space.

In a *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ we have the *set of outcomes* Ω, or *sample space*, the *set of events* $\mathcal{F}$ and the probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$ that assigns probabilities to events. Recall that $\Omega, \emptyset \in \mathcal{F}$ is always trivially so, hence the probability measure $\mathbb{P}$ is defined on all outcomes in the sample space Ω.

Example 1.7 (Discrete probability space)**.** Let $\Omega =$"a countable set", that is, finite or countably infinite. Take the σ-field $\mathcal{F} = 2^{\Omega}$ as the power set. Consider,

$$\mathbb{P}(A) = \sum_{\omega\in A} p(\omega) : p(\omega) \geq 0$$

and note

$$\sum_{\omega\in\Omega} p(\omega) = 1.$$

Hence, for a finite set Ω, we may define a *uniform* probability measure $p(\omega) = \frac{1}{|\Omega|}$. Such as for a fair six sided dice, where $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $p(\omega) = 1/6$. In particular, for some *event* $A \in \mathcal{F}$ where $A =$"we get a 2 and a 3", we have chance $\mathbb{P}(A) = 1/6 + 1/6 = 1/3$.

Theorem 1.8. *Suppose we are given some collection of σ-fields $\mathcal{F}_i, i \in I$, where the index set $I \neq \emptyset$ is arbitrary (i.e., possibly uncountable). Then $\cap_{i\in I}\mathcal{F}_i$ is a σ-field.*

Proof.

i.) Fix any $A \in \mathcal{F}_i$ for some $i \in I$. Then $A^c \in \mathcal{F}_i$ and so, if $A \in \cap_{i\in I}\mathcal{F}_i$ then $A^c \in \cap_{i\in I}\mathcal{F}_i$.

ii.) Take any countable sequence of sets $\{A_k\}_{k=1}^{n}$ then $\cup_{k=1}^{n} A_k \in \mathcal{F}_i$ and so, if $\{A_k\}_{k=1}^{n} \in \cap_{i\in I}\mathcal{F}_i$ then, $\cup_{k=1}^{n} A_k \in \cap_{i\in I}\mathcal{F}_i$.

□

2. Bernoulli Distribution.

$$\boxed{X \sim Bernoulli(p) : 0 \leq p \leq 1} \tag{1}$$

Take some random variable X with a Bernoulli parametric model with success parameter p defined by:

$$X \triangleq \begin{cases} 1 & \text{if success } p, \\ 0 & \text{otherwise.} \end{cases}$$

Then, $X \sim Bernoulli(p)$ read as; "X has distribution Bernoulli with parameter p". The probability mass function is then defined by:

$$\boxed{f_X(x; p) \triangleq p^x(1-p)^{1-x} : x \in \{0, 1\}.} \tag{2}$$

3. Binomial Distribution.

$$\boxed{X \sim Bin(n, p) : 0 \leq p \leq 1} \tag{3}$$

This is a finite generalisation of a series of Brnoulli random variables with replacement.

Take some random variable X with each observation $x_i \in X : 0 \leq i \leq n$ where n is the number of trials and that each $x_i \sim Bernoulli(p)$. Then, provided all the observations x_i are independent and since the parameter p is the same (a special case of the Poisson binomial distribution) for all. Then, we claim that the each x_i are "independent and identically distributed" **i.i.d**. Assuming now, we select each x_i with replacement, then there are $\binom{n}{x}$ ways to select x observations from a trial, or sample size of n.

That is, $X \sim Bin(n, p) : 0 \leq p \leq 1$ and hence has probability mass function defined by:

$$\boxed{f_X(x; n, p) \triangleq \binom{n}{x} p^x(1-p)^{n-x} : 0 \leq x \leq n.} \tag{4}$$

Clear, this is just an extension of a Bernoulli trial sequentially repeated n times.

4. Geometric Distribution.

$$\boxed{X \sim Geom(p) : 0 \leq p \leq 1} \tag{5}$$

This is a special case result of a Binomial random variable, for when $x = 1$.

Take some random variable X such that $X \sim Bin(n, p)$, given that p is the probability of success. Take $x = 1$ with $n = k$ so that, the probability of first success on the k^{th} trial is given by,

$$f_X(1; k, p) = \binom{k}{1} p^1(1-p)^{k-1}.$$

Then we have the following special case result as our probability mass function:

$$\boxed{f_X(k;p) \triangleq p(1-p)^{k-1} : k \in \{1,2,\dots\}.} \tag{6}$$

5. Poisson Distribution.

$$\boxed{X \sim Poisson(\lambda)} \tag{7}$$

This is another special case result of a Binomial random variable, for when p is close to zero and n is very large[Pap91].

Take some random variable Y such that $Y \sim Bin(n,p)$ and fix some parameter $\lambda \doteq np$ where n is known to be large and p is known to be small. Such a case is intuitively a rare event since the probability of success p is small for a large number of trials n. This is know as "the law of rare events"[Sem]. We continue with the assumtion that such events are sufficiently independent. Then, we have a limiting case of the Binomial random variable Y with parameter λ.

Hence if,

$$Y \sim Bin(n,p)$$

Substituting for λ and by introduction of an auxiliary random variable X

$$X \sim Bin(n, \frac{\lambda}{n})$$

we derive the probability mass function as follows,

$$\begin{aligned} f_X(x;\lambda) &= \frac{n!}{(n-x)!x!}(\frac{\lambda}{n})^x(1-\frac{\lambda}{n})^{n-x} \\ &= \frac{n!}{n^x(n-x)!}\frac{1}{(1-\frac{\lambda}{n})^x}\frac{\lambda^x}{x!}(1-\frac{\lambda}{n})^n \end{aligned}$$

Taking the limit as $n \to \infty$ and noting the definition of $e^{-\lambda} \doteq \lim_{n\to\infty}(1-\frac{\lambda}{n})^n$

$$\begin{aligned} &\frac{\lambda^x}{x!}\lim_{n\to\infty}\left(\frac{n!}{n^x(n-x)!}\right)\left(\frac{1}{(1-\frac{\lambda}{n})^x}\right)\left(1-\frac{\lambda}{n}\right)^n \\ &= \frac{\lambda^x}{x!}(1)(1)(e^{-\lambda}) \end{aligned}$$

hence, we have the probability mass function to be:

$$\Rightarrow f_X(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} : x \in \{0,1,2,\dots\} \text{ for some fixed } \lambda.$$

Thus, Given a random variable X such that $X \sim Poisson(\lambda)$ then X has probability mass function:

$$\boxed{f_X(x;\lambda) \triangleq \frac{e^{-\lambda}\lambda^x}{x!} : x \in \{0,1,2,\dots\}} \tag{8}$$

6. Hypergeometric Distribution.

$$\boxed{X \sim Hypergeom(M,N,n) : \max(0, M+n-N) \leq k \leq \min(M,n)} \tag{9}$$

Bernoulli trials without replacement (Binomial is actually a special case of the hypergeometric). Probability mass function is defined as:

$$\boxed{\begin{aligned} f_X(x;M,N,n) \triangleq \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} &: N \in \{1,2,\dots\} \\ & m \in \{0,1,2,\dots,N\} \\ & n \in \{1,2,\dots,N\} \end{aligned}} \tag{10}$$

References

[Pap91] Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, NY, USA, 1991.

[Sem] Ladislaus Semali. Home page.