Probability A

Jack Betteridge

22/6/14

Def$^{\underline{n}}$. Random Experiments: All outcomes form an outcome/sample space
Ω - the set of all outcomes
An outcome is some $\omega \in \Omega$
An event is some $A \subset \Omega$
$\emptyset$ is the impossible event.

Def$^{\underline{n}}$. Classical probability spaces with equally likely events:
A finite model $\Omega = \{1, 2, ..., n\}$, $\mathcal{P}(\omega) = \frac{1}{n}$, where $\omega \in \Omega$, but we want to define probability on events not outcomes

Def$^{\underline{n}}$. Let $\Omega \neq \emptyset$. A system of subsets of Ω, $\mathcal{A}$ is called an algebra if:

i) $\Omega \in \mathcal{A}$

ii) $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$

iii) $A \in \mathcal{A} \implies A^{\complement} \in \mathcal{A}$

Cor$^{\underline{ly}}$. $A \cap B \in \mathcal{A}$

Def$^{\underline{n}}$. $B_1, \ldots, B_n \subset \Omega$ is called a partition if:

i) $\bigcup_{i=1}^{n} B_i = \Omega$

ii) $B_i \cap B_j = \emptyset \quad \forall i \neq j$

Def$^{\underline{n}}$. Let $\Omega \neq \emptyset$ & $B_1, \ldots, B_n$ be a partition of Ω

1. $(\Omega, \mathcal{A})$ is called a classical measurable space & $b_1, \ldots, B_n$ are called basic events.
2. $(\Omega, \mathcal{A}, P)$ is called a classical probability space if:
   - $(\Omega, \mathcal{A})$ is a classical measureable space
   - $P : \mathcal{A} \to [0, 1]$ is a set function $s.t.$ $\forall A \in \mathcal{A}$ comprised of k basic events B_i, $P(A) = \frac{k}{n}$

Def$^{\underline{n}}$. A measurable space $(\Omega, \mathcal{A}^*)$ is called a submodel of $(\Omega, \mathcal{A})$ if $\mathcal{A}^* \subset \mathcal{A}$

Th$^{\underline{m}}$. *Given*

$$\begin{array}{lll} n_1 & \textit{distinct elements} & a_1, \ldots, a_{n_1} \\ n_2 & \textit{distinct elements} & b_1, \ldots, b_{n_2} \\ & \vdots & \\ n_\nu & \textit{distinct elements} & x_1, \ldots, x_{n_\nu} \end{array}$$

$\exists N = \prod_{i=1}^{\nu} n_i$ *distinct ordered ν-tuples* $(a_i, \ldots, x_{\nu_i})$

Def$^{\underline{n}}$. A premutation of a finite set is any ordering of its elements in a list.

Th$^{\underline{m}}$. *A set with n elements has $n!$ different permutations.*

Def$^{\underline{n}}$. Sampling Types: If we sample r from n distinct elements the method is called:

- with replacement if any element that was drawn is replaced *s.t.* it may be drawn again.
- without replacement if once drawn an element cannot be drawn again.

The record is called:

- ordered if the order is recorded, use: (, ,)
- unordered id the order is not recorded, use: { , , }

Th$^{\underline{m}}$. *Suppose we sample $r \geq 1$ from n distinct elements. Let $O(r,n)$ be the number of distinct ordered records & $N(r,n)$ be the number of distinct not ordered records. Then:*

- *with replacement:* $O(r,n) = n^r$, $N(r,n) = \binom{n+r-1}{r} = \frac{(n+r-1)!}{r!(n-1)!}$
- *without replacement:* $O(r,n) = \frac{n!}{(n-r)!}$, $N(r,n) = \binom{n}{r} = \frac{n!}{r!(n-r)!}$

Th$^{\underline{m}}$. *In allocating r indistinguishable elements to n coordinates, the number of distinct allocations N is:*

i) $\binom{n+r-1}{r}$

ii) if $r \geq n$ & no position unoccupied $\binom{r-1}{r-n}$

Def$^{\underline{n}}$. Let $\Omega \neq \emptyset$. A system of subsets of Ω, $\mathcal{A}$ is called a σ-algebra if:

i) $\Omega \in \mathcal{A}$

ii) $A_i \in \mathcal{A}$ $(i \in \mathbb{N}) \implies \bigcup_{i\in\mathbb{N}} B_i = \Omega$

iii) $A \in \mathcal{A} \implies A^{\complement} \in \mathcal{A}$

Def$^{\underline{n}}$. $B_i \subset \Omega$ $(i \in \mathbb{N})$ is a partition if it is exhaustive *i.e* : $\bigcup_{i\in\mathbb{N}} B_i = \Omega$ and mutually exclusive *i.e* : $B_i \cap B_j = \emptyset$ $\forall i \neq j$

Def$^{\underline{n}}$. Let $(\Omega, \mathcal{A})$ be a measurable space. A set function $P : \mathcal{A} \to [0,1]$ is called a probability measure if:

i) $P(\Omega) = 1$

ii) P is σ-additive: $\forall A_i \in \mathcal{A}$ $(i \in \mathbb{N})$ with $A_i \cap A_j = \emptyset$ $\forall i \neq j$ $P(\bigcup_{i\in\mathbb{N}}) = \sum_{i\in\mathbb{N}} P(A_i)$

$(\Omega, \mathcal{A}, P)$ is called a probability space.

Deductions from Axioms:

i) Empty sets: $P(\emptyset) = 0$

ii) Finite Additivity: $\forall A_i \in \mathcal{A}$ $(i \in \{1,\dots,n\})$ with $A_i \cap A_j = \emptyset$ $\forall i \neq j$, $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$

iii) Complement: $P(A^{\complement}) = 1 - P(A)$

iv) Partition by another event: $P(A) = P(A \cap B) + P(A \cap B^{\complement})$

v) Difference: $P(A \backslash B) = P(A) - P(A \cap B)$

vi) Subset: $B \subset A \implies P(A \backslash B) = P(A) - P(B)$

vii) Monotony: $B \subset A \implies P(B) \leq P(A)$

viii) Addition Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

ix) Subadditivity: $P(\bigcup_{i\in\mathbb{N}} A_i) \leq \sum_{i\in\mathbb{N}} P(A_i)$

Def$^{\underline{n}}$. The smallest σ-algebra containing all open intervals (hence all closed intervals & all countable unions and intersections of open and closed intervals) is called a Borel σ-algebra and on $\mathbb{R}$ is denoted $\mathcal{B}(\mathbb{R})$.

Prop$^{\underline{n}}$. *If $\Omega = [0,1],\ \omega \in \Omega,\ P(\{\omega\}) = 0$, then for random numbers on $[0,1]$, $P([a,b)) = b - a$*

Def$^{\underline{n}}$. Let P be a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. If $F(x) = P((-\infty, x])$ for $x \in \mathbb{R}$ satisfies:

i) F is non-decreasing $i.e : x > y \implies F(x) \geq F(y)$

ii) $\lim_{x \searrow -\infty} F(x) = 0$ & $\lim_{x \nearrow \infty} F(x) = 1$

iii) F is continuous from the right $i.e : \forall x_0 \in \mathbb{R}\ \lim_{x \searrow x_0} F(x) = F(x_0)$

then it is called a cumulative distribution function for P.

Prop$^{\underline{n}}$. *There is a one to one correspondence between probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and cumulative distribution functions.*

Def$^{\underline{n}}$. Probability distributions can be classified as:

i) Discrete: $P(\{x_0\}) > 0$, x_0 is an atom, P has finitely many or countably finitely many atoms & F is a step function.

ii) Continuous F is differentiable & called density.

iii) Mixed: A combination of (i) & (ii).

Def$^{\underline{n}}$. Let $(\Omega, \mathcal{A}, P)$ be a probability space & $A, B \in \mathcal{A}$. Given $P(B) > 0$. The conditional probability of A given B is $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

Th$^{\underline{m}}$. *Given the above $P(\cdot|B) : \mathcal{A} \to [0,1]$ is a probability space?????*

Th$^{\underline{m}}$.
Multiplication Rule: $P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$
"Flip around" Formula: $P(A|B) = \frac{P(A)}{P(B)} \cdot P(A|B)$
Averaging Conditional Probabilities: $P(A) = P(A|B) \cdot P(B) + P(A|B^{\complement}) \cdot P(B^{\complement})$
Baye's Theorem: $P(B|A) = \frac{P(A|B)P(B)}{P(A|B) \cdot P(B) + P(A|B^{\complement}) \cdot (1 - P(B))}$

Th$^{\underline{m}}$. *Let $(\Omega, \mathcal{A}, P)$ be a probability space & $B_1, \dots, B_n$ be a partition of Ω,*
then $P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$.

Th$^{\underline{m}}$. *(Baye's Theorem) Let $(\Omega, \mathcal{A}, P)$ be a probability space & $B_1, \dots, B_n \in \mathcal{A}$ a partition of Ω, then*
$P(B_k|A) = \frac{P(A|B_k) \cdot P(B_k)}{\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)}$.
$P(B_i)$ are called prior-probabilities known before A.
$P(A|B_i)$ are called likelihoods - probabilities of A given different B_i.
$P(B_i|A)$ are called posterior-probabilities after knowing about A.

Def$^{\underline{n}}$. Let $(\Omega, \mathcal{A}, P)$ be a probability space & $A, B \in \mathcal{A}$. A & B are called independent if $P(A \cap B) = P(A) \cdot P(B)$ and dependent otherwise.

Th$^{\underline{m}}$. *General multiplication rule:* $P(A_1 \cap \cdots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot \cdots \cdot P(A_n|A_1 \cap \cdots \cap A_{n-1})$.

Def$^{\underline{n}}$. Events are pairwise independent if $P(A_j \cap A_k) = P(A_j) \cdot P(A_k) \quad \forall j, k \in \{1, \dots n\}, j \neq k$

Bernoulli Trials:

i) 2 possible outcomes: Success or Failure.

ii) Probability $p \in [0,1]$ for success is the same for each experiment.

$q = 1 - p$ - probability for failure.
$\Omega = \{0,1\}^n$ or $\Omega = \{0,1\}^{\mathbb{N}}$
$S_k =$"k successes in n trials"$= \{\omega : \sum_{i=1}^{n} \omega(i) = k\}$ where $k \in \{0, \dots, n\}$
$P(\{\omega\}) = p^k q^{n-k} \quad \forall \omega \in S_k$ hence $P(S_k) = \binom{n}{k} p^k q^{n-k}$

Def$^{\underline{n}}$. The probability measure on the measure space $(\{0, \dots, n\}, \mathcal{P}(\{0, \dots, n\}))$ with weights $p_k = \binom{n}{k} p^n q^{n-k}$ for $k \in \{0, \dots, n\}$ is called the binomial distribution with parameters n & p, denoted $\text{Bin}(n, p)$.

Def$^{\underline{n}}$. The probability measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ with weights $p_k = q^{k-1} p$ with $k \in \{1, \dots\}$ is called the geometric distribution with parameter p, denoted $\text{Geom}(p)$.

Def$^{\underline{n}}$. The probability measure on $(\{r, r+1, \dots\}, \mathcal{P}(\{r, r+1, \dots\}))$ with weights $p_k = \binom{k-1}{r-1} p^r q^{k-r}$ with $k \in \{r, r+1, \dots\}$ is called the negative binomial distribution with parameters p & r.

Th$^{\underline{m}}$. *Let $\lambda > 0$ & $(p_n)_{n \in \mathbb{N}}$ where $p_n \in [0,1] \quad \forall n$ with $np_n \to \lambda$, then $\forall k \in \{0, 1, \dots\}$, $\lim_{n\to\infty} P(H_k(n, p_n)) = e^{-\lambda} \frac{\lambda^k}{k!}$, where $H_k(n,p) = Bin(n,p)$ with k successes.*

Def$^{\underline{n}}$. Let $\mu > 0$. The probability measure on $(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0))$ with weights $p_k = e^{-\lambda} \frac{\mu^k}{k!}$ is called the Poisson distribution with parameter μ, denoted $\text{Po}(\mu)$.

Th$^{\underline{m}}$. *Using $Bin(\infty, p)$ & $S_n(\omega) = \sum_{i=1}^{n} \omega_i =$ "N$^{\underline{o}}$ successes" $\forall \epsilon > 0$ $P(|\frac{1}{n} S_n - p| < \epsilon) \to 1$ as $n \to \infty$.*