Dr. Osonde Osoba

This Week:

- Review
  - Sequences, series, infinite series summations
  - Counting measure, Binomial distribution
- Random Variables
  - Definition and Illustration. $P \circ X^{-1}(A)$
  - Properties: Independence, measurability vs Continuity
- Examples of Random variables: Discrete (BEG CUP)
  - Binomial, Hypergeometric, Geometric/NB
  - Poisson
  - Approximation Laws.
    - Poisson Law
    - Hypergeometric

H/W 3:

L-G: 2.44 – 2.47, 2.58 – 2.61

I/ Review:

Given $\{a_k\}_{k\in\mathbb{N}}$ [sequence]

we defined $S_n = \sum_{k=0}^{n} a_k$ [series]

Then the infinite series S is:

$$\boxed{S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{k=0}^{n} a_k}$$

This definition informs our discussions of P-measures [eg CA]

Recall: $\lim_{k\to\infty} a_k = L$

$\iff \forall \varepsilon>0,\ \exists n_0\in\mathbb{N} : \forall k\geq n_0 \quad |a_k - L| < \varepsilon$

$\iff$ all but a finite number of $\{a_k\}_k$ are outside

any ε-neighbourhood of L.

We will use these to discuss measures on discrete sample spaces. For example: Consider the finite countable product space

$$\Omega = \{0,1\}^n \qquad [\sim \{H,T\}^n]$$

If every element $\vec{\omega}\in\Omega$ is "equally likely," we use the counting measure on the σ-algebra $\mathcal{Q} = 2^{\Omega}$:

$$\boxed{P(\{\vec{\omega}\}) = \frac{1}{|\Omega|}} = 2^{-n}$$

Suppose instead that the relative frequency of occurrence for $\vec{\omega}_k$ depends on the contents of the sequence $\vec{\omega}_k$.

e.g. in a coin flip $\vec{\omega}_k = (H, T, T, H, T)$

$$P(\{\vec{\omega}_k\}) = \prod_{i=1}^{n} p^{\omega_{k,i}} (1-p)^{1-\omega_{k,i}} \quad [\equiv a_k]$$

where $p = P(\{\omega_{k,i} = H\}) \in [0,1]$

The Binomial Theorem guarantees that the sum of all 2^n such a_k is 1 [T].

since $\sum_{k=1}^{2^n} a_k = (p + (1-p))^n = 1^n$

The Binomial (measure) distribution is concerned with Events, E_k, of the form:

$$E_k = \{\vec{\omega} \in \Omega \mid \vec{\omega} \text{ has } k \text{ Hs and } (n-k) \text{ Ts}\}$$

Thus we know $|E_k| = \binom{n}{k}$

and

$$\boxed{P(E_k) = \binom{n}{k} a_k = \binom{n}{k} p^k (1-p)^{n-k}}$$

Since $E_k = \bigcup_{\vec{\omega} \in E_k} \{\vec{\omega}\}$ and $P(\{\vec{\omega}\}) = a_k \;\forall \vec{\omega} \in E$

<u>Aside</u>: We can define a smaller σ-algebra $\mathcal{Q}_1 \subset 2^\Omega$ made up of only such E_k.

Qu: Are all outcomes equally likely in a sample space $\Omega = \{H,T\}^n$ for a fair coin?

Note:

① We have relied on the absolutely summable sequence $\{a_k = \binom{n}{k} p^k (1-p)^{n-k}\}_{k=0}^{n}$ for the def^n of $P(\cdot)$ on this discrete space.

② This is a cumbersome description of a probability space/measure. Random variables simplify things a bit...

II Random Variables:

ⓐ A random variable is a function

$$X : \Omega \rightarrow \mathbb{R}$$

such that $X^{-1}(A) \in \mathcal{Q} \quad \forall A \in \mathcal{B}(\mathbb{R})$

ⓑ A random variable, X, is a measurable map from the measurable space $(\Omega, \mathcal{Q})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

A measurable map between $(\Omega, \mathcal{Q})$ and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is any function $X : \Omega_X \rightarrow \mathbb{R}$ such that $X^{-1}(A) \in \mathcal{Q} \quad \forall A \in \mathcal{B}(\mathbb{R})$

rewrite

eg: (i) suppose $A \in \mathcal{Q}$ [Indicator function $1_A(\omega)$]

$$X : (\Omega, \mathcal{Q}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

$$X(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

$$\Rightarrow X^{-1}(\{1\}) = A \in \mathcal{Q}, \quad X^{-1}(\{0\}) = A^c \in \mathcal{Q}$$

$\therefore$ X is an r.v.

$$\left[\forall B \in \mathcal{B}(\mathbb{R}), \; X^{-1}(B) \in \begin{Bmatrix} X^{-1}(\{1\}), & X^{-1}(\{0\}) \\ X^{-1}(\phi), & X^{-1}(\{0,1\}) \end{Bmatrix} \right]$$

(ii) $\Omega = \{H, T\}^n, \quad \mathcal{Q} = 2^{\Omega}$

$$\tilde{\Omega} = \{0, 1, 2, \cdots, n\}, \quad \mathcal{F} = 2^{\tilde{\Omega}}$$

$$X : (\Omega, \mathcal{Q}) \longrightarrow (\tilde{\Omega}, 2^{\tilde{\Omega}})$$

$$X(\vec{\omega}) = \#(H_s \text{ in } \vec{\omega}) \quad \left[\sum_{i=1}^{n} \delta(\omega_i = H) \right]$$

$\mathcal{Q}$ and $\mathcal{F}$ are power sets. So.... X is an r.v.

And

$$P(X = k) \triangleq P(X^{-1}(\{k\})) = P \circ X^{-1}(\{k\})$$

$$= P(\{\omega \in \Omega \mid X(\omega) = k\})$$

$$= P(E_k) = \binom{n}{k} p^k (1-p)^k$$

Aside

rewrite

We talk about the p-measure induced by the r.v. $X : (\Omega, \mathcal{Q}) \to (S, \mathcal{F})$ as the p-measure on the range measurable space $(S, \mathcal{F})$ such that

$$\boxed{\forall A \in \mathcal{F} \qquad P_X(A) = P(X^{-1}(A))}$$

i.e. $$P_X(A) = P(\{\omega \in \Omega \mid X(\omega) \in A\})$$

we also write this as $\boxed{P(X \in A)}$

When $(S, \mathcal{F}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $A \in \mathcal{B}(\mathbb{R}^d)$ we call $P(X \in A) = P_X(A)$ the distribution of X.

Recall that $X^{-1}(\cdot)$ preserves (arbitrary) unions, intersections, complementation, and even subsethood. So we can convert all probability concerns from $(\Omega, \mathcal{Q}, P)$ to $(S, \mathcal{F}, P_X)$ without issue.

$\underline{\sigma(X)}$: The r.v. X also induces a σ-algebra on Ω

$$\boxed{\sigma(X) = \{X^{-1}(B) \subset \Omega \mid B \in \mathcal{F}\}}$$

This is the preimage σ-algebra.

Claim: $\sigma(X)$ is a σ-algebra

Proof: $X : \Omega \longrightarrow (S, \mathcal{F})$

(i) $X^{-1}(S) = \Omega$ [defn of a fxn]

$S \in \mathcal{F}$ [$\mathcal{F}$ is a σ-alg]

$\Rightarrow \Omega \in \sigma(X)$ $\therefore$ [T]

(ii) $B \in \mathcal{F} \Rightarrow B^c \in \mathcal{F}$

$X^{-1}(B) \in \sigma(X)$

$(X^{-1}(B))^c = X^{-1}(B^c) \in \sigma(X)$ [$\because B^c \in \mathcal{F}$]

$\therefore$ [C]

(iii) $\{B_i\}_i \subset \mathcal{F} \Rightarrow X^{-1}(B_i) \in \sigma(X)$ $\forall i$

$\bigcup_i X^{-1}(B_i) = X^{-1}(\bigcup_i B_i)$ [Property : X^{-1}]

$\bigcup_i B_i \in \mathcal{F}$ [defn : σ-alg]

$\therefore X^{-1}(\bigcup_i B_i) \in \sigma(X)$

$\therefore$ [U]

$\sigma(X)$ is CUT $\therefore$ $\sigma(X)$ is a σ-alg.

$\sigma(X)$ is the smallest σ-alg on Ω that makes $X : \Omega \to S$ a measurable map.

rewrite

The use of random variables $X : (\Omega, \mathcal{Q}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ allows us to ignore the fine structure of Ω and focus just on $P_X(\cdot)$, the distribution induced by X.

Recall (from homework problems) that the sets $A = \{k\}$ are also Borel sets. We can write the distribution of discrete random variables X (i.e $X(\omega) \in \mathbb{Z}\ \forall\, \omega$) in terms of $P(X(\omega) \in \{k\})$ or $P(X = k)$. This is the probability distribution function or probability mass function $p_X(k)$

Ex: i) $X \sim \text{Bernoulli}(p) \qquad X \in \{0, 1\}$

$$p_X(k) = \begin{cases} p & k = 1 \\ (1-p) & k = 0 \end{cases}$$

ii) $X \sim \text{Binomial}(n, p) \qquad X \in \{0, 1, \cdots, n\}$

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

iii) $X \sim \text{Hypergeometric}(n, N_1, N_2)$

$$P_X(k) = \frac{\binom{N_1}{k}\binom{N_2}{n-k}}{\binom{N_1+N_2}{n}}$$

Aside

Claim: $\binom{N_1+N_2}{n} = \sum_{k=0}^{n} \binom{N_1}{k}\binom{N_2}{n-k}$

Proof: [Vandermonde Identity]

$$(1+x)^{N_1+N_2} = \sum_{n=0}^{N_1+N_2} \binom{N_1+N_2}{n} x^n \quad \ldots (*)$$

$$= (1+x)^{N_1} \cdot (1+x)^{N_2}$$

$$= \left(\sum_{k=0}^{N_1} \binom{N_1}{k} x^k\right)\cdot\left(\sum_{j=0}^{N_2} \binom{N_2}{j} x^j\right)$$

$$= \sum_{n=0}^{N_1+N_2} \left(\sum_{k=0}^{n} \binom{N_1}{k}\binom{N_2}{n-k} x^n\right) \quad \ldots (+)$$

equate coefficients of x^n in $(*)$ and $(+)$

$$\therefore \binom{N_1+N_2}{n} = \sum_{k=0}^{n} \binom{N_1}{k}\binom{N_2}{n-k}$$

iv) $X \sim \text{Uniform}(1, \ldots, m)$

$$P_X(k) = 1/m \quad \forall\, k \in \{1, \ldots, m\}$$

Qu: $E_k = \{\text{Coin flipped until } k^{th} \text{ time before first } H\}$

$P(E_k) = ?$ ~~IA~~ $[P(H) = p]$

Ans:

$E_k \Rightarrow (k-1)$ failures and then 1 success

$$P(E_k) = (1-p)^{k-1} \cdot p$$

vi $X \sim \text{Geometric}(p) \qquad X \in \{1, \dots\} \qquad p \in (0,1)$

$\{X = k\} = \{\text{repeat independent Bernoulli}(p) \text{ until } 1^{st} \text{ success @} k\}$

$$p_X(k) = (1-p)^{k-1} \cdot p$$

Qu: Show $\sum_k p_X(k) = 1 \quad [X \sim \text{Geometric}(p)]$

Ans recall by Binomial Theorem [generalized]

$$(1-p)^{-1} = \sum_{k=0}^{\infty} p^k \qquad |p| < 1$$

$$\Rightarrow p(1-p)^{-1} = \sum_{k=1}^{\infty} p^k$$

$$\sum p_X(k) = p \sum_{k=1}^{\infty} (1-p)^{k-1} = p \cdot \sum_{k=0}^{\infty} q^k$$

$$= p \cdot (1-q)^{-1}$$

$$= p/(1-(1-p)) = p/p$$

$$= 1$$

Aside:

Measurability vs. Continuity

There is a duality between σ-algebras and topologies: $\mathcal{Q} \subseteq 2^{\Omega}$, $\mathcal{T} \subseteq 2^{\Omega}$

σ-algebra, $\mathcal{Q}$	Topology, $\mathcal{T}$
i/ $\Omega \in \mathcal{Q}$	i/ $\Omega \in \mathcal{Q}$, $\phi \in \mathcal{Q}$
ii/ $A \in \mathcal{Q} \Rightarrow A^c \in \mathcal{Q}$	ii/ $\{A_t\}_{t\in T} \subset \mathcal{T} \Rightarrow \left(\bigcup_{t\in T} A_t\right) \in \mathcal{T}$
iii/ $\{B_i\}_{i\in I} \subset \mathcal{Q} \Rightarrow \bigcup_{i\in I} B_i \in \mathcal{Q}$	iii/ $\{E_s\}_{s\in S} \subset \mathcal{T} \Rightarrow \left(\bigcap_{s\in S} E_s\right) \in \mathcal{T}$
[I countable]	[T abitrary, S finite]

$A \in \mathcal{Q} \Leftrightarrow A$ is a measurable set

$E \in \mathcal{T} \Leftrightarrow E$ is an open set

$(\Omega, \mathcal{Q})$: measurable space

$(\Omega, \mathcal{T})$: topological space

$X : (\Omega, \mathcal{Q}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a measurable function

$\Leftrightarrow X^{-1}(B) \in \mathcal{Q} \quad \forall B \in \mathcal{B}(\mathbb{R})$ //

$X : (\Omega, \mathcal{T}) \longrightarrow (\mathbb{R}, \mathcal{T}_{\mathbb{R}})$ is a continuous function

$\Leftrightarrow X^{-1}(B) \in \mathcal{T} \quad \forall B \in \mathcal{T}_{\mathbb{R}}$ //

A group of pdfs (BEG CUP) cover many typical probabilistic senarios. The combinatorial pdfs (Binomial, Hypergeometric, Geometric/NB) fall under the B of BEG CUP:

# of Outcomes	Sampling scheme: with replacement	Sampling scheme: w/o replacement
$n=2$	Binomial "Until": – Geometric – Negative Binomial	Hypergeometric
$n \geq 3$	Multinomial	Multivariate Hypergeometric

B – Binomial [hyp., multinomial, NB/Geom.]

E – Exponential [Gamma, n-Erlang, $\chi_i^2(r)$]

G – Gaussian

C – Cauchy [special case of $S\alpha S$]

U – Uniform [Continuous and discrete, Beta generalization]

P – Poisson

Binomial Approximation for Hypergeometric pdfs:

Given a box of m black balls and $N-m$ red balls, mixed together. Sample n balls without replacement. X = # of black balls in n.

$\Rightarrow \quad X \sim Hyp(n, m, N-m)$

The effect of sampling w/o replacement diminishes as the total number of balls increases. $N \to \infty$

This suggest a simpler approximation for $P(X=k)$ as $N \to \infty$: Hypergeometric $\xrightarrow{d}$ Binomial$(N,\ p = m/N)$

$$\lim_{N\to\infty} P(X=k) = \lim_{N\to\infty} \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$$

$$= \lim_{N\to\infty} \left(\frac{m!}{k!\,(m-k)!}\right)\left(\frac{(N-n)!\,n!}{N!}\right)\left(\frac{(N-m)!}{(n-k)!\,(N-m-n+k)!}\right)$$

$$= \lim_{N\to\infty} \frac{n!}{(n-k)!\,k!} \cdot \frac{\dfrac{m!\,(N-m)!}{(N-m-n+k)!\,(m-k)!}}{\dfrac{N!}{(N-n)!}}$$

$$= \lim_{N\to\infty} \binom{n}{k} \frac{\overbrace{\dfrac{m!}{(m-k)!}}^{(m)_k} \cdot \overbrace{\dfrac{(N-m)!}{(N-m-n+k)!}}^{(N-m)_{n-k}}}{\dfrac{N!}{(N-n)!}}$$

$$\frac{m!}{(m-k)!} = m \times (m-1) \times \cdots \times (m-k+1) = \prod_{j=1}^{k} (m-k+j)$$

$$\frac{(N-m)!}{(N-m-n+k)!} = \prod_{j=1}^{n-k} (N-m-n+k+j)$$

$$N^n = N^k \cdot N^{n-k}$$

$$\Rightarrow \lim_{N\to\infty} P(X=k) = \lim_{N} \frac{\prod_{j=1}^{k} (m-k+j) \cdot \prod_{j=1}^{n-k} (N-m-n+k+j)}{\prod_{j=1}^{n} (N-n+j)} \cdot \frac{N^n}{N^n} \cdot \binom{n}{k}$$

$$= \binom{n}{k} \cdot \lim_{N\to\infty} \frac{\left[N^k \prod_{j=1}^{k} (m-k+j)\right]\left[N^{n-k} \prod_{j=1}^{n-k} (N-m-n+k+j)\right]}{\left[N^n \prod_{j=1}^{n} (N-n+j)\right]}$$

$$= \binom{n}{k} \lim_{N\to\infty} \frac{\prod_{j=1}^{k} \left(\frac{m}{N} \underbrace{- \frac{k}{N} + \frac{j}{N}}_{\to 0}\right) \cdot \prod_{j=1}^{n-k} \left(1 - \frac{m}{N} \underbrace{- \frac{n}{N} + \frac{k}{N} + \frac{j}{N}}_{\to 0}\right)}{\prod_{j=1}^{n} \left(1 \underbrace{- n/N + j/N}_{\to 0}\right)}$$

$$= \binom{n}{k} \left(\frac{m}{N}\right)^k \left(1 - m/N\right)^{n-k}$$

$$= \binom{n}{k} p^k (1-p)^{n-k}$$

$$= b(n, k, p)$$

The Poisson Distribution :

Recall :

$$e^x = \sum_{n=0}^{\infty} x^n/n! \quad \ldots \quad (1)$$

$$e^x = \lim_{n\to\infty} \left(1 + x/n\right)^n \quad \ldots \quad (2)$$

(1) ... is an absolutely convergent series valid $\forall x \in \mathbb{R}$

Use this to define the Poisson distribution:

$$X \sim \text{Poisson}(\lambda) \iff \boxed{P_X(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}} \quad \begin{array}{l} k \geq 0 \\ \lambda \in \mathbb{R}^+ \end{array}$$

$$\Rightarrow \sum_{k=0}^{\infty} P_X(k) = e^{-\lambda} \sum_{k=0}^{\infty} \lambda^k/k! = e^{-\lambda} \cdot e^{\lambda} = 1 \,//$$

$P(X = k)$ describes the probability of observing k events in a fixed unit interval of time/length/volume/etc. When the average # of events per unit time is λ.

Poisson distributions apply to random phenomena where 3 conditions hold roughly for $\Delta t << 1$:

i) $P(1 \text{ event} \in (t, t+\Delta t)) = \lambda \cdot \Delta t + o(\Delta t)$ $[\forall t]$

ii) $P(k \text{ events} \in (t, t+\Delta t)) = o(\Delta t)$ $k > 1$

iii) $P(n \text{ events in } I_n \text{ and } m \text{ events in } I_m) = P(n \text{ events in } I_n) \cdot P(m \text{ events in } I_m)$ if $I_n \cap I_m = \emptyset$

The Poisson is also the limiting distribution of the Binomial as $n \gg 1$ and $p \ll 1$ with $\lambda = np$ constant

$\approx$ n independent intervals with Binary events in each interval

<u>Theorem:</u> $\quad b(n, p) \xrightarrow[p \ll 1]{n \gg 1} P(\lambda = np)$

<u>Proof:</u>

$$\lim_{n} \binom{n}{k} p^k (1-p)^{n-k} = \lim_{n} \frac{(n)_k}{k!} p^k (1-p)^{n-k}$$

$$= \lim_{n} \frac{(n)_k}{k!} \left(\lambda/n\right)^k \left(1-\lambda/n\right)^{n-k}$$

$$\lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} = \left(\lambda^k/k!\right) \cdot \lim_{n} \left[\frac{(n)_k}{n^k} \cdot \left(1-\lambda/n\right)^{n-k}\right]$$

$$(n)_k / n^k = \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{(n-k+1)}{n} \longrightarrow 1 \text{ as } n \to \infty$$

$$\lim_{n\to\infty} \left(1-\lambda/n\right)^{-k} = 1$$

$$\Rightarrow \lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} = \lambda^k/k! \cdot \left[\lim_{n\to\infty} \left(1-\lambda/n\right)^n\right] \cdot 1$$

$$= \lambda^k/k! \cdot e^{-\lambda}$$

Ex: 10E4 components with independent annual probability of failure 10E-3. ~~in a year.~~ System fails if any component fails. P(system running 1 year later) = ?

$$P(\text{running}) = P(0 \text{ failures out of } n = 10E4 \text{ tries})$$

$$= \binom{10e4}{0} (10e{-3})^0 (1 - 10e{-3})^{10e4}$$

$$= (1 - 10e{-3})^{10e4}$$

Poisson approx: $\lambda = np = 10$; $X \sim \text{Poisson}(10)$ ← # component failures in year

$$P(\text{running}) \approx P(X=0) = e^{-10} \cdot (10)^0/0! = e^{-10}$$