STA347 Probability - Lecture notes

Heather Tan

Sep - Dec 2019

Contents

1 List of Common Sets

- Natural Number: $\mathbb{N} = \{1, 2, \cdots\}$
- Whole: $\mathbb{W} = \{\mathbb{N}\} \cup \{0\}$
- Integers: $\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$
- Rationals: $\mathbb{Q} = \{\frac{n}{m} | n \in \mathbb{Z}, m \in \mathbb{N}\}$
- Reals: $\mathbb{R} = \{x = \lim_{n\to\infty} r_n | r_n \in \mathbb{Q}, n \in \mathbb{N}\}$
- Complex: $\mathbb{C} = \{z = x + iy | x, y \in \mathbb{R}\}$

2 Sequence of Event A_n

2.1 sigma-additivity

Definition 2.1. P is said to be $\sigma-$*additive/countably additive* *iff* for any mutually disjoint sequence of events $A_n, n \in N$, $P(\sum_1^\infty A_n) = \sum_1^\infty P(A_n)$

2.2 Continuity

Definition 2.2. $A_n \to A \implies P(A_n) \to P(A)$

2.3 Indicator Function

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

2.3.1 One-to-one

$$\begin{aligned} I_A = I_B &\implies I_A(\omega) = I_B(\omega) \forall \omega \\ &\implies \omega \in A \iff I_A(\omega) = I_B(\omega) = 1 \\ &\iff \omega \in B \implies A = B \end{aligned}$$

2.3.2 Onto

Take any $f \in \{0,1\}^\Omega$ and simply let $A = \{\omega | f(\omega) = 1\}$ then by the very definition of this set

$$\begin{aligned} &\omega \in A \iff f(\omega) = 1 \iff I_A(\omega) = 1 \implies f(\omega) = I_A(\omega) \forall \omega \\ &f = I_A \end{aligned}$$

2.4 Infinitely Often

$$\omega \in \Omega, A_n \subset \Omega, n \in \mathbb{N}$$

$$\omega \in \cup_{n=1}^{\infty} A_n, \omega \in \cap_{n=1}^{\infty}$$

$$\exists\ ns.t\omega \in A_n \iff \omega \in \cap_{n=1}^{\infty} A_n$$

There is at least one A_n $s.t$ $\omega \in A_n$

$$P(\omega \in A_n, \exists n) = P(\omega \in \cup_{n=1}^{\infty} A_n)$$

$$P(\omega \in A_n, \forall n) = P(\omega \in \cap_{n=1}^{\infty} A_n)$$

$$P(\omega \in A_n, io(n)) = P(\omega \in \cap_{N=1}^{\infty} \cup_{n=N}^{\infty} A_n)$$

Infinitely Often: $\forall N, \exists n \geq N$ $s.t$ $\omega \in \cup_{n=N}^{\infty} A_n \iff \omega \in \cap_{N=1}^{\infty} \cup_{n=N}^{\infty} A_n)$

Alternatively, recall indicates functions:

$$A \subset \Omega \text{ denote } I_A : \Omega \to \{0,1\} = 2$$

$$\text{by} I_A(\omega) = \begin{cases} 1, \omega \in A \\ 0, \omega \notin A \end{cases}$$

Now define the indicator map: $I : p(\Omega) \to \{0,1\}^{\Omega}$, $A^B = \{f : B \to A\}$, $A \mapsto I_A$

Note

$$I_{\cap_{n=N}^{\infty} A_n}(\omega) = inf_{n=1}^{\infty} I_{A_n}(\omega)$$

$$I_{\cup_{n=N}^{\infty} A_n}(\omega) = sup_{n=1}^{\infty} I_{A_n}(\omega)$$

$$\sum_{n=1}^{\infty} I_{A_n}(\omega) \in \mathbb{N} \cup \{0\} \cup \{\infty\} = \mathbb{W} \cup \{\infty\}$$

$$\forall N, \exists n \geq N \ s.t\ \omega \in \cup_{n=N}^{\infty} A_n \iff \omega \in \cap_{N=1}^{\infty} \cup_{n=N}^{\infty} A_n \iff \sum_{n=1}^{\infty} I_{A_n}(\omega) = 0$$

2.5 Convergence

Definition 2.3. $A_n \to A \iff I_{A_n} \to I_A$ io $I_{A-n}(\omega) \to I_A(\omega) \forall \omega \in \Omega$

Recall the meaning of $x_n \to x \implies$ cauchy $sup_{i,j \geq n} |x_i - x_j| \to 0$ as $n \to \infty = sup_{i \geq n} x_i - inf_{j \geq n} x_j$

For: $inf_{j \geq n} x_j \leq x_n, x \leq sup_{i \geq n} x_i$

- $m \leq n \implies inf_{j \geq m} x_j \leq inf_{j \geq n} x_j \leq x_n \leq sup_{i \geq n} x_i$
- $m \geq n \implies inf_{j \geq m} x_j \leq x_n \leq sup_{j \geq m} x_i \leq sup_{i \geq n} x_i$

$\forall m, n \quad inf_{j \geq m} x_j \leq sup_{i \geq n} xi$

Theorem 2.1. $x_n \to x \iff inf_{j\geq n}x_j \to x \quad and \quad sup_{i\geq n}x_j \to x$

Definition 2.4. $\underline{lim}x_n \overset{name}{=} lim_{n\to\infty}infx_n \overset{defn}{=} lim_{n\to\infty}inf_{j\geq n}x_j = sup_{n=1}^{\infty}inf_{j=n}^{\infty}x_j$

Definition 2.5. $\overline{lim}x_n \overset{name}{=} lim_{n\to\infty}supx_n \overset{defn}{=} lim_{n\to\infty}inf_{i\geq n}x_i = inf_{n=1}^{\infty}sup_{i=n}^{\infty}x_j$

$inf_{j\geq m}x_j \leq \underline{lim}\ x_n \leq \overline{lim}\ x_n \leq sup_{i\geq n}x_i$

- $I(\cup_{n=1}^{\infty}A_n) = sup_{n=1}^{\infty}I(A_n)$
- $I(\cap_{n=1}^{\infty}A_n) = inf_{n=1}^{\infty}I(A_n)$

2.5.1 Theoretic limits

$$\begin{aligned} I_{A_n} \to I_A \implies lim_{n\to\infty}I(A_n) = \underline{lim}_{n\to\infty}I_{A_n} &= sup_{n=1}^{\infty}inf_{i=n}^{\infty}I(A_k) \\ &= I(\cup_{n=1}^{\infty}\cap_{k=n}^{\infty}A_k) \\ = \overline{lim}_{n\to\infty}I_{A_n} &= inf_{n=1}^{\infty}sup_{i=n}^{\infty}I(A_k) \\ &= I(\cap_{n=1}^{\infty}\cup_{k=n}^{\infty}A_k) \end{aligned}$$

Be sure to able to understand this case:

$$\begin{aligned} &(A_t, t\in T) \\ \omega \in \cup_{t\in T}A \iff \omega &\in A_t \exists t \in T \\ I_{\cup_{t\in T}A_t} &= sup_{t\in T}I_{A_t}(\omega) \end{aligned}$$

$$\begin{aligned} &I_{A_n} \to I_A \\ &\iff I(\cup_{n=1}^{\infty}\cap_{i=n}^{\infty}A_i) = I(\cap_{n=1}^{\infty}\cup_{i=n}^{\infty}A_i) \\ &\iff \cup_{n=1}^{\infty}\cap_{i=n}^{\infty}A_i = \cap_{n=1}^{\infty}\cup_{i=n}^{\infty}A_i = A \end{aligned}$$

$$A_n \to A \iff \cup_{n=1}^{\infty}\cap_{i=n}^{\infty}A_i = \cap_{n=1}^{\infty}\cup_{i=n}^{\infty}A_i$$

2.5.2 Set-theoretic terms

$A_n \to A \iff A = \cup_{n=1}^{\infty}\cap_{k=n}^{\infty}A_k = \cap_{n=1}^{\infty}\cup_{k=n}^{\infty}A_k$

2.5.3 Notation of Convergence Sequence

- increasing sequence of sets: $A_n \subset A_{n+1}$, then the limit of this sequence not only exists but is simply the union of the sets:$A_n \uparrow \implies \cup_{n=1}^{\infty}A_n$
- decreasing sequence of sets: $A_{n+1} \supset A_n$, then the limit of this sequence is: $A_n \downarrow \cap_{n=1}^{\infty}A_n$

3 FTAP - Fundamental Theorem of Applied Probability

Theorem 3.1. *$U = \sum_{i=1}^{\infty} p^{-i} Z_i$, based on $p \geq 2$, we find that $U \sim unif\ [0,1] \iff Z_i$ iid $unif\ \{0, 1, \cdots, p-1\}$ generate uniform from 0 to 1.*

Proof.

$$\begin{aligned}
P &= \{0, 1, \cdots, p-1\} \\
\mathbb{R}^n &= \{x \mid x_i \in \mathbb{R}, i = 1, \cdots, n\} \\
x &= (x_i, i = 1, \cdots, n) \\
\mathbb{R}^\infty &= \{x = (x_i, i = 1, \cdots, n) \mid x_i \in \mathbb{R}, i \in \mathbb{N}\}
\end{aligned}$$

Lemma: let $p^\infty = \{\underline{x} \mid x_i \in p, i \in \mathbb{N}\}$ $\quad \dot{p}^\infty = \{\underline{x} \mid x_i \in p, i \in \mathbb{N}, \text{but not allowed to end in p-1 repeat}\}$
$U = \sum_{i=1}^{\infty} p^{-i} Z_i$ defines a 1-1 and onto function $\Phi : \dot{p}^\infty \to [0, 1)$ $\square$

Note: $\sum_{i=1}^{\infty} a^i \overset{defn}{=} lim_{n\to\infty} \sum_{i=1}^{n} a^i = lim_{n\to\infty} \frac{a^{n+1}-1}{a-1} = lim_{n\to\infty} \frac{1-a^{n+1}}{1-a} = \frac{1}{1-a}$

Definition 3.1. $Z \sim unif\ (p), p = \{0, 1, \cdots, p-1\} \iff p(z = i) = 1/p \quad \forall i \in p$

$\dot{p}^\infty = \{\underline{z} = (z_i, i \in \mathbb{N}) \mid z_i \in p = \{0, 1, \cdots, p-1\}, i \in \mathbb{N}, z_i < p-1, io(i)\}$

Lemma 3.1. $u = \sum_{i=1}^{\infty} z_i p^{-i} \quad \underline{z} = (z_i, i \in \mathbb{N}) \in \dot{p}^\infty \quad \implies z_i = b_1, \cdots, z_n = b_n$
$\iff u \in [\ \sum_{i=1}^{n} b_i p^{-i}, \sum_{i=1}^{n} b_i p^{-i} + p^{-n})$

Lemma 3.2. *$u = \sum_{i=1}^{\infty} z_i p^{-i} \quad \underline{z} = (z_i, i \in \mathbb{N}) \in \dot{p}^\infty$ defines a bijection: $\Phi : \dot{p}^\infty \overset{\approx}{\to} [0, 1)$, i.e $\underline{z} \mapsto u$*

Proof. $0 \leq u < (p-1) \sum_{i}^{\infty} p^{-i} = \frac{p-1}{p} \sum_{o}^{\infty} p^{-i} = \frac{p-1}{p} \frac{1}{1-1/p} = 1$
thus

$$\begin{aligned}
u = \sum_{1}^{\infty} z_i p^{-i} \implies 0 \leq u - \sum_{1}^{n} z_i p^{-i} &= \sum_{n+1}^{\infty} z_i p^{-i} < (\sum_{n+1}^{\infty} z_i p^{-i})(p-1) \\
&= p^{-(n+1)} \frac{1}{1-1/p} (p-1) \\
&= p^{-(n+1)} \frac{p}{p-1} (p-1) \\
&= p^{-n}
\end{aligned}$$

$$\begin{aligned}
&\iff 0 \leq u - \sum_{1}^{n} z_i p^{-i} < p^{-n} \\
&\iff z_n p^{-n} \leq u - \sum_{1}^{n-1} z_i p^{-i} < (z_n + 1) p^{-n} \\
&\iff z_n \leq p^n (u - \sum_{1}^{n-1} z_i p^{-i}) < (z_n + 1)
\end{aligned}$$

[x] = m iff $m \leq x < m+1$ uniquely determine m as the greatest integer less than or equal to x

$$\iff z_n = p^n(u - \sum_{1}^{n-1} z_i p^{-i}) \quad n \geq 2$$

$$z_1 = [pu]$$

□

Definition 3.2. $Z \sim unif\{0, \cdots, p-1\}, p(Z \in \{0, \cdots, p-1\}) = 1 \iff P(Z = i) = P(Z = j) \forall i, j = 0, \cdots, p-1$

Theorem 3.2. $U = \sum_{i=1}^{\infty} Z_i p^{-i} \implies U \sim unif[0,1] \impliedby Z_i IIDZ \sim unif\{0, \cdots, p-1\}$

If p=10, $U = \sum_{i=1}^{\infty} Z_i p^{-i} = Z_1 Z_2 Z_3$ which can be considered as digits of decimal numbers.

4 Probability Space $(\Omega, \mathcal{F}, \mathcal{P})$

- The sample space Ω – an arbitrary non-empty set
- The σ-algebra $\mathcal{F} \subseteq 2^{\Omega}$(also called σ-field) – a set of subsets of Ω, called events, s.t:
  - $\Omega \in \mathcal{F}$
  - $\mathcal{F}$ is closed under complements: $A \in \mathcal{F} \implies (\Omega / A) \in \mathcal{F}$
- The probability measure $P : \mathcal{F} \to [0,1]$ – a function of $\mathcal{F}$ s.t
  - P is countably additive(σ-additive) if $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ is a countable collection of pairwise disjoint sets, then $P(\cup_{i=1}^{\infty} A_i) = \sigma_{i=1}^{\infty} P(A_i)$
  - $P(\Omega) = 1$

Definition 4.1. $a \in \mathcal{P}(\Omega) \iff A \subset \Omega$

$$\begin{aligned} \Omega &= set, \\ \emptyset \neq \mathcal{F} &\subset \mathcal{P}(\Omega), \\ A_1, A_2, \cdots, A_n, \cdots \in \mathcal{F} \implies \cap_1^{\infty} A_n &\in \mathcal{F} \\ A &\in \mathcal{F}, \\ A^c &\in \mathcal{F} \end{aligned}$$

Proof. $\cap_{i=}^{\infty} A_i = (\cup_{i=}^{\infty} A^c)^c$
$A \subset \mathcal{F}, A \cap A^c = \emptyset \in \mathcal{F}$ □

4.1 Sigma-Addictivity

$P(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$, P have to be non-neg. σ additive set functions: $\mu(\cup_{n=1}^{N} A_n) = \sum_{n=1}^{N} \mu(A_n)$.
$P : \mathcal{F} \overset{Prob}{\to} \mathbb{R}$ have to be

- sigma-add
- normed
- non-neg

Theorem 4.1. $A_n \to A \implies P(A_n) \to P(A)$

5 Same in Distribution $\overset{d}{=}$

5.1 Samples equal in distribution

$x \overset{d}{=} y$ on sample space $\mathscr{X} \iff Eg(x) = Eg(y) \quad \forall g : \mathscr{X} \to \mathbb{R} \iff g = I_A, A \subset \mathscr{X}$
$P(X \in A) = EI_A(X) = EI_A(Y) = P(Y \in A)$

Theorem 5.1. $X \overset{d}{=} Y \implies \phi(X) \overset{d}{=} \phi(y)$ *for any* $\phi : \mathscr{X} \to \mathscr{Y}$

Proof. $Eh\phi(X) = Eh\phi(y) \quad \forall h : (Y) \to \mathbb{R}$ ☐

5.2

$E : \mathscr{R} \to \mathbb{R} \cup \{\pm\infty\} \cup \{*\}$ (the star aka undefine) by the very way in which E is defined on $\mathscr{R} = \{X :$ real-valued ramdon variable$\}$

Sample Questions

$$\begin{aligned}
((n+1) - w)^3 &\overset{d}{=} w^3 \\
(n+1)^3 - 3(n+1)^2 + 3(n+1)w^2 - w^3 &\overset{d}{=} w^3 \\
\implies 2EW^3 = (n+1)^3 - 3(n+1)^2 EW + 3(n+1)^2 EW^2 \\
\implies EW^3 = n(EW)^2 \\
EW^2 - E(W-1)^2 = n \\
\implies EW^2 = \frac{2n+1}{3} EW
\end{aligned}$$

Homework

- EW^4?
- $EW^5 - E(W-1)^5 = n^4$

6 Expected Value, variance and Covariance

Definition 6.1. $EX \overset{defn}{=} lim_{n\to\infty} \frac{x_1+\cdots+x_n}{n}$

Definition 6.2. $EX^2 = \sum_1^n x_i^2/n$

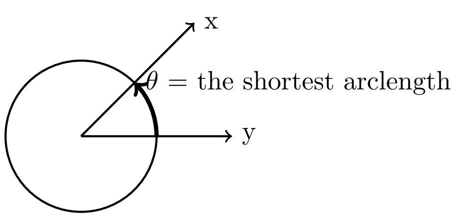
Definition 6.3. $EXY = \sum_1^n x_i y)i/n$

6.1 Length of Vectors

$||X|| = \sqrt{EX^2} = \lim_{n\to\infty} \frac{\sqrt{\sum_1^n x_i^2}}{\sqrt{n}} = lim_{n\to\infty} \frac{|x|}{\sqrt{n}}$
$\langle X, Y\rangle = EXY = lim_{n\to\infty} \underline{x}'y/n = \sum_1^n x_i y_i$

6.2 Measured Angle

$\measuredangle(\underline{x}, y) = \theta(\underline{x}, y) =$ angle between x and y



$$cos\theta(\underline{x}, y) = \frac{\underline{x}'\underline{y}}{|x||y|} = \frac{\underline{x}'y/n}{(|\underline{x}|\sqrt{n})(|\underline{y}|\sqrt{n})} \to \frac{EXY}{\sqrt{EX^2}\sqrt{EY^2}}$$

6.2.1 Connection with r

$\dot{X} = X - EX \quad \dot{Y} = Y - EY$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}} \overset{conv}{\to} \frac{E\dot{X}\dot{Y}}{\sqrt{E\dot{X}^2}\sqrt{E\dot{Y}^2}} = \frac{E(X-EX)(Y-EY)}{||X-EX||\ ||Y-EY||} = cos\measuredangle(\dot{X}, \dot{Y})$$

Definition 6.4. $cos\angle(x, y) \overset{defn}{=} \frac{EXY}{||X||\ ||Y||} \overset{LLN}{=} \frac{\underline{x}'\underline{y}}{|x||y|} = cos\measuredangle(x, y)$

Definition 6.5. $\rho(x, y) = cos\measuredangle(\dot{X}, \dot{Y}) = \frac{E(X-EX)(Y-EY)}{\sqrt{varX}\sqrt{varY}} = \frac{cov(x,y)}{\sigma(x)\sigma(y)}, \quad \sigma(x) = \sqrt{varX} = ||\dot{x}|| = ||X - EX||$

6.3 Expected Value

Proposition 6.1. *EX is the closest constant to X:* $||X - EX|| = inf_{t\in\mathbb{R}}||x - t||$

Proof. hint: f(t) = $\sqrt{E(x-t)^2}, \quad g(t) = E(X-t)^2$, try to minimize g(t) $\square$

6.3.1 properties of $E(\dot{X})$

- $(X \dot{+} Y) = \dot{X} + \dot{Y} = X - EX + Y - EY$
  $X + Y - E(X+Y) = X + Y - EX - EY$
- $(c\dot{X}) = cX - EcX = c(X - EX) = c\dot{X}$
- $E\dot{X} = 0$

6.3.2 Expected value for arbitrary finite discrete distribution(Lebesgue-stieltjes)

$X \sim \begin{pmatrix} a_1, \cdots, a_n \\ p_1, \cdots, p_n \end{pmatrix}$
$\iff x = \sum_{j=1}^N a_j I_{\{a_j\}}(x) \implies EX = \sum_{j=1}^N a_j EI_{\{a_j\}}(x) = \sum_{j=1}^N a_j P(X = a_j) = \sum_{j=1}^N a_j p_j$

consider any $\mathbb{R}$-value $x \geq 0$ and let

$$\begin{aligned} 0 \leq x_n &= \sum_{j=1}^n \frac{j-1}{\sqrt{n}} I_{(\frac{j-1}{\sqrt{n}}, \frac{j}{\sqrt{n}}]}(x) \leq x \\ 0 \leq x - x_n &= \sum_{j=1}^n (x - \frac{j-1}{\sqrt{n}}) I_{(\frac{j-1}{\sqrt{n}}, \frac{j}{\sqrt{n}}]}(x) + xI_{(\sqrt{n},\infty)}(x) \\ &\leq \sum_{j=1}^n \frac{1}{\sqrt{n}} I_{(\frac{j-1}{\sqrt{n}}, \frac{j}{\sqrt{n}}]}(x) + xI_{(\sqrt{n},\infty)}(x) \\ &= \frac{1}{\sqrt{n}} I_{(0,\sqrt{n}]}(x) + xI_{(\sqrt{n},\infty)}(x) \\ &\leq \frac{1}{\sqrt{n}} + xI_{(\sqrt{n},\infty)}(x) \to 0 \text{ as } n \to \infty \end{aligned}$$

and so we find $x_n \to x$ as $n \to \infty$
But now let $h : [0, \infty) \overset{cont}{[}0, \infty$(i.e continuous map to real number of non- negative) and we automatically have $h(x_n) \to h(x) \implies Eh(x_n) \to Eh(x)$
i.e

$$\begin{aligned} Eh(x_n) &= lim_{n\to\infty} h(\frac{j-1}{\sqrt{n}}) p(\frac{j-1}{\sqrt{n}} < x < \frac{j}{\sqrt{n}}) \\ &= lim_{n\to\infty} h(\frac{j-1}{\sqrt{n}}) [F(\frac{j}{\sqrt{n}}) - F(\frac{j-1}{\sqrt{n}})] \\ &\overset{name}{=} \int_0^\infty h(x) dF(x) \\ x \geq 0 \implies xI_{(\sqrt{n},\infty)}(x) &= x[\sum_{j=1}^n I_{(\frac{j-1}{\sqrt{n}}, \frac{j}{\sqrt{n}}]}(x) + xI_{(\sqrt{n},\infty)}(x)] \\ &= \sum_{j=1}^n xI_{(\frac{j-1}{\sqrt{n}}, \frac{j}{\sqrt{n}}]}(x) + xI_{(\sqrt{n},\infty)}(x) \end{aligned}$$

6.4 Covariance

6.4.1 Basic Calculation

- $cov(x, y) \overset{defn}{=} E\dot{x}\dot{Y} = E(X - EX)(Y - EY)$
- $cov(x, x) = EX^2 = E(X - EX)^2 = var(X)$
- $cov(X + Y, Z) = cov(X, Z) + cov(Y, Z)$
- $cov(cX, Z) = c\ cov(X, Z)$

6.4.2 Three Properties: Bi-linear,Symmetric, non-negative

- bi-linear: $cov(\sum_1^m a_i x_i, \sum_1^n b_j y_j) = \sum_i^m \sum_j^n a_i b_j cov(x_i, y_i)$
- symmetrci: $cov(x, y) = cov(y, x)$
- non-negative definition: $cov(x, x) = var(x) \geq 0$ w.eq(with equality) $\iff X \overset{defn}{=} EX$

6.5 Markov Inequality and Chebychev Inequality

Proposition 6.2. *given $Z \geq 0, t \geq 0$, any $g : [0, \infty) \to [0, \infty)$ with $g \uparrow \implies P(Z \geq t) \leq \frac{Eg(Z)}{g(t)} \forall g$ that satisfies all above hypothesis*

Proof. $g(Z) \geq g(t)I(Z \geq t) \implies Eg(Z) \geq EI(Z \geq t) = g(t)P(Z \geq t)$ $\square$

Corollary 6.1. $P(|X - EX| \geq k) \leq \frac{E(X-EX)^2}{k^2}$ *i.e* $P(|X - EX| \geq k\sigma) \leq \frac{1}{k^2}$

application: $varX = 0 \iff x \overset{wp}{=} EX$

Proof.
$(\implies)$
$varX = 0 \implies P(|X - EX| \geq \frac{1}{n}) \leq n^2 \times 0 = 0$
i.e $P(|X - EX| \geq \frac{1}{n}) = 0 \quad \forall n$
$P(|X - EX| < \frac{1}{n}) = 1 \quad \forall n$
$A_n = (|X - EX| < \frac{1}{n}) \downarrow$ as $n \to \infty$
$A = (|X - EX| = 0) \implies P(A_n) \to P(A)$, so $P(A) = 1$
$P(|X - EX| = 0) = 1$
$|x| = 0 \iff x = 0$
equivalently $P(X = EX) = 1$ and so we write $X \overset{wp1}{=} EX$ which is merely a symbol for the above.
$(\impliedby)$
$P(X = EX) = 1 = P(|X - EX|^2 = 0)$ so let any $y = (X - EX)^2 = P(y = 0) = 1$
we find $varX = Ey = 0 \times 1 = 0$ $\square$

6.6 Conditional Expectation Probability

Define:

- $E(x|w) \overset{defn}{=} o.p(X|L_2(w))$ with o.p refers to orthogonal projection and x is R-valued
- $L_2(w) = \{g(w)|g : \mathcal{W} \to \mathbb{R}, Eg(w)^2 < w\}$
- $\mathcal{W}$ = sample space of W
- $\mathcal{R}$: all R-valued random variable
- L_2 : all R-valued random variable x s.t $Ex^2 < \infty$(Counter Example: take $Z \sim N(0,1), andletx = \frac{1}{z}, x^2 = \frac{1}{z^2}$)

Now note $\mathcal{R}$ itself is a vector space.

Example: find $cov(\sum_{i=1}^m a_i x_i, \sum_{j=1}^n b_j y_j) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j cov(x_i, y_i)$, given $x, y \in L_2$ i.e $Ex^2 < \infty, Ey^2 < \infty \implies E|XY| < \infty$

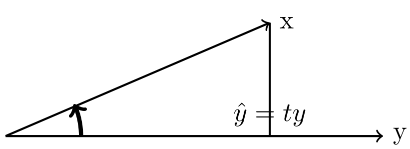
Proof. $|XY| < \frac{X^2+Y^2}{2}$ □

so $EXY \in \mathbb{R}$

Corollary 6.2. $X, Y \in L_2 \implies X + Y \in L_2$

i.e L_2 is a sub-vector space

Proof. $(X+Y)^2 = X^2 + 2XY + Y^2 \leq X^2 + 2|XY| + Y^2$ □



Definition 6.6. o.p $x|y = \hat{y}$

This satisfices two properties:

- $\hat{y} = ty$ for some $t \in \mathbb{R}$
- $x - \hat{y} \perp y \implies E((x-\hat{y})y) = 0 \iff E(x-ty)y = 0 \iff t = \frac{EXY}{EY^2}$

7 Probability Mass Function(PMF)

$$F(x) = P(X \leq x) = P(X \in (\infty, x)) = P_x(-\infty, x]$$
$$F(x + n^{-1}) = P_x(-\infty, x + n^{-1}] \to F(x) = P_x(-\infty, x]$$

we say that any distribution function is not continuous at any point x

Definition 7.1. $F(x+) = lim_{n\to\infty} F(x + \frac{1}{n}) = F(x)$

Definition 7.2. $F(x-) = lim_{x\to\infty} F(x - \frac{1}{n}) = P(X < x)$

Definition 7.3. the probability math function of X is $p : \mathbb{R} \to [0, 1]$ given by $p(x) = P(X = x)$

Proposition 7.1. $p(x) = F(x) - F(x-)$

Proof. $p(x) = P(X \leq x) - P(X < x) = F(x) - F(x-)$ □

Proposition 7.2. $^{\#}(p > 0) \leq^{\#} N$

Proof.

$$(p > 0) = \cup_{n=1}^{\infty}(p > \frac{1}{n})$$
$$(p > 0) = p^{-1}(0, \infty) = \{x \in \mathbb{R}|p(x) > 0\}$$

but $\#(p > \frac{1}{n}) < n$ because o.w $\exists a_1, \cdots, a_n \in (p > \frac{1}{n})$, so if we let $A = \{a_1, \cdots, a_n\} \implies$ then we find $P(x \in A) = \sum_{i=1}^{n} P(x = a_i) = \sum_{i=1}^{n} p(a_i) > \frac{n}{n} = 1$ □

e.g let $\mathbb{Q} = \{r_n|n \in \mathbb{N}\}$, let $\{p_n|n \in \mathbb{N}\}$ when $p_n > 0$ s.t $\sum_{n\in\mathbb{N}} p_n = 1$. We simply define $F(x) = \sum_{\{n|r_n \leq x\}} p_n$

8 Uniform Distribution

Definition 8.1. $U \sim Unif(\Omega),^{\#} \Omega <^{\#} \mathbb{N} \iff P(U = \omega) = \frac{1}{\#\Omega} \iff P(U \in A) = \frac{\#A}{\#\Omega}$

e.g

$$U \sim unif\ \{1, \cdots, n\} \iff P(U = i) = \frac{1}{n}, i = 1, \cdots, n$$

$$\begin{aligned}\text{note that:} \quad -U &\sim unif\ \{-n, \cdots, -1\}\\ n - U &\sim unif\ \{0, 1, \cdots, n-1\}\\ n - U + 1 &\sim unif\ \{1, \cdots, n\}\end{aligned}$$

$$\left.\begin{aligned} n + 1 - U &\overset{d}{=} U\\ n + 1 - EU &= EU\end{aligned}\right\} \implies EU = \frac{n+1}{2}$$

Definition 8.2. $U \sim unif[0, 1] \iff P(U \leq u) = u, \forall\ 0 \leq u \leq 1$

8.1 $U \sim unif\ \{1, \cdots, n\}$, $EU^k - E(U - 1)^k = n^{k-1}$

- Since $U \overset{d}{=} n + 1 - U \quad EU = E(n + 1 - U) = n + 1 - EU \implies EU = \frac{n+1}{2} = \frac{1+\cdots+n}{n}$
- $U \sim unif\{1, \cdots, n\} \quad U - 1 \sim unif\{0, \cdots, n\} \implies$ U and U-1 has the similar distribution that differ by a shift.

- $U^k \sim unif\ \{1^k, 2^k, \cdots, n^k\}$ $\quad EU^k = \frac{1+2^k+\cdots+n^k}{n} \implies E(U-1)^k = \frac{0+1+2^k+\cdots+(n-1)^k}{n}$
- $EU^k = \frac{1}{k+1} \quad k \in$

8.2 Variance of U and its calculation

Result: $Var(U) = Var(U-1) = Var(n+1-U)\frac{n^2-1}{12}$

$$
\begin{aligned}
&EU^3 - E(U-1)^3 = n^2 \\
&\implies EU^3 - (EU^3 - 3EU^2 + 3EU - 1) = n^2 \\
&\implies 3EU^2 = n^2 + 3EU - 1 = (n-1)(n+1) + \frac{3(n+1)}{2} = \frac{(n+1)(2n+1)}{2} \\
&EU^2 = \frac{(n+1)(2n+1)}{6} = \frac{1+2^2+\cdots+n^2}{n} \\
\text{recall: } &Var(U) = EU^2 - (EU)^2 = \frac{(n+1)(2n+1)}{6} - \frac{n+1}{2}\frac{n+1}{2} \\
&= (\frac{2n+1}{3} - \frac{n+1}{2})\frac{n+1}{2} = \frac{n-1}{6}\frac{n+1}{2} = \frac{n^2-1}{12} \\
&\implies Var(U) = Var(U-1) = Var(n+1-U)\frac{n^2-1}{12} \text{ Since } (var(aX+b) = a^2 varX)
\end{aligned}
$$

8.3 U and nU

8.3.1 $U \sim unif[0,1] \iff [nU] \sim unif\{0, \cdots, n-1\} \forall n$

Proof.

$(\implies)$

$P([nU] = k) = P(k \le nU \le k+1) = \frac{1}{n}$ provided $k = 0, 1, \cdots, n-1$

$(\impliedby)$

to show that $P(U \le u) = u \quad \forall 0 \le u \le 1$, consider $P(U > r)$ for any $r \in \mathbb{Q}[0,1)$

But $P(U < r) = P(nU < k)$ for $r = k/n, k < n$

$= \sum_{i=0}^{n-1} P(nU < k, [nU] = i) = \sum_{i=0}^{n-1} P(0 \le nU < k, i \le nU < i+1)$

i.e $\cap P$ on $[0, k]$ which is the accumulation of P(nU) in each partition

$= \sum_{i=0}^{k-1} P(i \le nU < i+1) = \sum_{i=1}^{k-1} P([nU] = i)$

$= k \cdot \frac{1}{n} = \frac{k}{n} = r$

$\implies P(U < r) = r$ i.e $P(U \le u)$ any other $0 \le u < 1$

Take $r_n \downarrow u$ at $r_n < 1 \quad [0, r_n) \downarrow [0, U]$

$P(U < r_n) = r_n \to P(U < u) = u$ $\square$

8.3.2 $U_n \to U$

$U \sim unif[0,1] \iff nU \sim unif\{0,\cdots,n-1\} \forall n \in \mathbb{N} \implies 0 \le U - \frac{nU}{n} \le \frac{1}{n}$
so $U_n = \frac{[nU]}{n} \to U$

9 Bernoulli Distribution

Definition 9.1. $Z \sim bern(p) \iff Z \sim \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$, $Z^{-1} \sim \begin{pmatrix} \infty & 1 \\ q & p \end{pmatrix}$

note: $Z^s = Z, x > 0$

10 Binomial Distribution

Definition 10.1. $X \sim bin(n,p), n \in \mathbb{N}, o \le p \le 1 \iff x \overset{d}{=} z_1 + \cdots + z_n, z_i iid, Z \sim bern(p)$

10.1 Properties

- $EX = \sum_1^n EZ_i = nEZ = np$
- $VarX = var(Z_1,\cdots,Z_n) \overset{indep}{=} VarZ_1,\cdots,VarZ_n + 0\cdot cov(Z_i,Z_j) \overset{ident}{=} nVarZ = npq$
- $\sigma(x) = \sqrt{n}\sqrt{pq}$

Proposition 10.1. *$X \sim bin(m,p), Y \sim bin(n,p)$ (independent) $\implies X + Y \sim bin(n+m,p)$*

Proof. $Z_1,\cdots,Z_{m+n}$ *idd*, $Z \sim bern(p) \implies \begin{pmatrix} X \\ Y \end{pmatrix} \overset{d}{=} \begin{pmatrix} Z_1 + \cdots + Z_m \\ Z_{m+1} + \cdots + Z_{m+n} \end{pmatrix} \implies X + Y \overset{d}{=} \sum_1^{m+n} Z_i$ □

e.g $g(Z_1,\cdots,Z_n) = \sum_{i=1}^n a_i Z_i^i \implies Eg(Z) = (\sum_{i=1}^n a_i)p \quad varg(Z) = (\sum a_i^2)pq$

10.2 Probability mass function

$p(k) = P_k = P(X=k) = P(Z_1 + \cdots, +Z_n = k)$ but $P(Z=z) = p^{-z}q^{1-z}, z \in \{0,1\}$

$$\begin{aligned}
\implies P((Z_1,\cdots,Z_n) = (z_1,\cdots,z_n)) &= P(Z_1 = z_1,\cdots,Z_n = z_n) \\
&\overset{indep}{=} P(Z_1 = z_1)\cdots P(Z_n = z_n) = p^{z_1}q^{1-z_1}\cdots p^{z_n}q^{1-z_n} \\
&= p^{\sum z_i} q^{n - \sum z_i} \\
\implies P(X=k) = \sum P((Z_1,\cdots,Z_n) &= (z_1,\cdots,z_n)) \quad z_1,\cdots z_n \in C_k^n
\end{aligned}$$

where $C_k^n = \{(z_1, \cdots, z_n) | z_i \in \{0,1\}, i = 1, \cdots, n \ s.t \ \sum_1^n z_i = k\}$
thus

$$\begin{aligned}
P(X = k) &= \sum_{(z_1,\cdots,z_n)\in C_k^n} p^{\sum_1^n Z_i} q^{n-\sum_1^n Z_i} \\
&= \sum_{(z_1,\cdots,z_n)\in C_k^n} p^k q^{n-k} \\
&= p^k q^{n-k} \times^{\#} C_k^n \\
&= \binom{n}{k} p^k q^{n-k}
\end{aligned}$$

10.2.1 $\binom{n}{k}$

$\binom{n}{k} \overset{name\#}{=} C_k^n$
$\sum_{k=0}^n P(X = k) = 1 = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k}$
let $p = \frac{t}{1+t}, q = \frac{1}{1+t}$
to find $(1+t)^n = \sum_{k=0}^n \binom{n}{k} t^k$

$$\begin{aligned}
n(1+t)^{n-1} &= \sum_{k=0}^n k\binom{n}{k} t^{k-1} \\
\overset{also}{=} n\sum_{k=0}^{n-1}\binom{n-1}{k} t^k &= n\sum_{j=1}^n \binom{n-1}{j-1} t^{j-1} \\
&= \sum_{j=1}^n n\binom{n-1}{j-1} t^{j-1} \\
&= \sum_{k=1}^n n\binom{n-1}{k-1} t^{k-1}
\end{aligned}$$

Theorem 10.1. $\sum_0^n a_k z^k = \sum_0^n b_k z^k \quad z_0 - \delta < z < z_0 + \delta \iff a_k = b_k \forall k = 0, \cdots, n$

$$\begin{aligned}
\implies k\binom{n}{k} &= n\binom{n-1}{k-1} = \frac{n^{(k)}}{k!} \\
\binom{n}{k} &= \frac{n}{k}\binom{n-1}{k-1} = \frac{n(n-1)}{k(k-1)}\binom{n-2}{k-2} \\
&= \frac{n(n-1)\cdots(n-(k-1))}{k(k-1)\cdots(k(k-1))}\binom{n-k}{0} = \frac{n^{(k)}}{k!} = \frac{n!}{k!(n-k)!}
\end{aligned}$$

10.2.2 properties of $n^{(k)}$

- $n^{(k)} = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}$
- $n^{(0)} = 1 = n^0 \quad 0! = 1$

e.g

$$\begin{aligned}
P(X=0) &= \binom{n}{0} p^0 q^n \\
&= p(Z_1+\cdots+Z_n=0) = P(Z_1=\cdots=Z_n=0) \\
&\overset{ind}{=} \prod_{i=1}^{n} p(Z_i=0) \overset{ind}{=} P(Z=0)^n = q^n
\end{aligned}$$

10.3 Expected Value

Proposition 10.2. $EX^{(k)} = \begin{cases} n^{(r)}p^r & r=0,\cdots,n \\ 0 & r=n+1,n+2 \end{cases}$

Proof. $g(x) = X^{(r)} = x(x-1)\cdots(x-r+1)$ polynomial of degree k, where $x \sim bin(n,p)$

$$\begin{aligned}
P(X=k) &= \binom{n}{k} p^k q^{n-k}, k=0,1,\cdots,n \\
EX^{(r)} &= Eg(x) = \sum_{k=0}^{n} g(k) \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=0}^{n} k^{(r)} \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=r}^{n} \frac{k(k-1)\cdots(k-(r-1))n^{(k)}}{k!} p^k q^{n-k} \\
&= \sum_{k=r}^{n} \frac{n^{(k)}}{(k-r)!} p^k q^{n-k} \\
&= \sum_{k=r}^{n} \frac{n^{(r)}(n-r)!p^r}{(k-r)![(n-r)-(k-r)]!} p^{k-r} q^{(n-r)-(k-r)} \\
&= n^{(r)} p^r \sum_{k=r}^{n} \binom{n-r}{k-r} p^{k-r} q^{(n-r)-(k-r)} \\
1 &= \sum_{k=r}^{n} \binom{n-r}{k-r} p^{k-r} q^{(n-r)-(k-r)} \\
\implies EX^{(r)} &= n^{(r)} p^r
\end{aligned}$$

□

10.4 Variance

- $EZ^2 = EZ = p$
- $VarZ = EZ^2 - E(Z)^2 = p - p^2 = pq$

11 Negative Binomial & Geometric Distribution

11.1 Definition

Given

- $Z \sim bern(p) \equiv \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$
- $Z_i, i \in \mathbb{N} \quad IIDZ$
- $S_n = \sum_{i=1}^{n} Z_i, n \in \mathbb{N} \quad S_{n+1} = S_n + Z_{n+1}$

Definition 11.1. $x \sim bern(n,p) \iff x \overset{d}{=} S_n$

Definition 11.2. $T_k \sim negbin(k,p) \iff (T_k = n) = (S_{n-k} = k-1, Z_n = 1), n = k, k+1, \cdots$

Interpretation of T_k: number of trials need to see k^{th} success.

Definition 11.3. $W \sim geo(p) \iff W \overset{d}{=} T_1, k-1 \implies geo(p) \equiv negbin(1,p)$

Note: $P(W=n) = P(T_1 = n) = P(S_{n-1} = 0, Z_n = 1) = P(S_{n-1} = 0)P(Z_n = 1) = pq^{n-1}, n = 1, 2, \cdots$

Proposition 11.1. $T \overset{d}{=} T_k, W \overset{d}{=} T_1, T \perp W \implies T + W \overset{d}{=} T_{k+1}$

Proof.

$$
\begin{aligned}
P(T+W=n) &= \sum_{i=1}^{n-k} P(T=n-i, w-i) \text{ since } n-i \geq k, i \leq n-k \\
&= \sum_{i=1}^{n-k} P(T=n-i)P(W=i) \\
&= \sum_{i=1}^{n-k} P(T_k=n-i)P(T_1=i) \\
&= \sum_{i=1}^{n-k} P(S_{n-i-1}=k-1, Z_{n-i}=1)P(S_{i-1}=0, Z_i=1) \\
&= \sum_{i=1}^{n-k} P(Z_1+\cdots+Z_{n-i-1}=k-1, Z_{n-i}=1)^{*}P(Z_{n-i+1}=\cdots=Z_{n-1}=0, Z_n=1) \\
&= \sum_{i=1}^{n-k} P(Z_1+\cdots+Z_{n-i+1}=k-1, Z_{n-i}=1, Z_{n-i+1}=\cdots=Z_{n-1}=0)P(Z_n=1)
\end{aligned}
$$

Let $j = n-i, i = n-j$,

$1 \leq i \leq n-k, -(n-k) \leq -i \leq -1$

$k \leq n-i \leq n-1, k \leq j \leq n-1$

$$= \sum_{j=k}^{n-1} P(Z_1 + \cdots + Z_{j-1} = k-1, Z_j = 1, Z_{j+1} = \cdots = Z_{n-1} = 0)P(Z_n = 1)$$

$$P(Z_1 + \cdots + Z_{n-1} = k) = \binom{n-1}{k} p^k q^{n-1-k}$$

$$= \sum_{j=k}^{n-1} P(Z_1 + \cdots + Z_n = k, \text{kth success occurs on jth trail})$$

$$= \sum_{j=k}^{n-1} P(Z_1 + \cdots + Z_{n-1} = k, Z_1 + \cdots + Z_{j-1} = k-1, Z_j = 1)$$

$$= \sum_{j=k}^{n-1} P(Z_1 + \cdots + Z_{j-1} = k-1, Z_j = 1, Z_{j+1} + \cdots + Z_{n-1} = 0)$$

$$\implies \sum_{n-1}^{j=k} P(Z_1 + \cdots + Z_{j-1} = k-1, Z_j = 1, Z_{j+1} = \cdots = Z_{n-1} = 0)P(Z_n = 1)$$

$$= P(Z_1 + \cdots + Z_{n-1} = k)P(Z_n = 1)$$

$$= P(S_{n-1} = k, Z_n = 1) \overset{defn}{=} P(T_{k+1} = n)$$

$\square$

Application:

$$T \overset{d}{=} T, W \overset{d}{=} T, T \perp W$$

$$\implies T + W \overset{d}{=} T_2$$

$$T \overset{d}{=} T, W \overset{d}{=} T, V \overset{d}{=} T \quad (T, W, V) statind.$$

$$T + W + V \overset{d}{=} (T + W) + V$$

Thus we see the obvious

Corollary 11.1. $W_1, \cdots, W_k \ IID, \ W \sim geo(p) \implies W_1 + \cdots + W_k \overset{d}{=} T_k$

Corollary 11.2. $T_1 \overset{d}{=} T_{k1}, T_2 \overset{d}{=} T_{k2}, T_{k1} \perp T_{k2} \implies T_1 + T_2 \overset{d}{=} T_{k1} + T_{k2}$

11.2 Expected value of T_k, S_n and W

Remember: $T_k \sim negbin(k, p) \iff (T_k = n) = (S_{n-1} = k-1, Z_n = 1) = (S_{n-1} < k < S_n)$

S_n = random number of successes in a fixed number n of trails.

T_k = random number of trails for a fixed number of k successes.

Note:

$$ES_n = np; \quad p = \frac{ES_n}{n}$$

p = average number of success per trail.

$$ET_k = \frac{k}{p}, \quad \frac{1}{p} = \frac{ET_k}{k}$$

$\frac{1}{p}$ = average number of trail per success

$$\begin{aligned} T_k &\overset{d}{=} \sum_{1}^{k} W_i, \quad W_i IID, \quad ET_k = kEW \\ EW &= \sum_{n=1}^{\infty} np(n) = \sum_{n=1}^{\infty} npq^{n-1} = p\sum_{n=1}^{\infty} nq^{n-1} \\ &= p\frac{d}{dq}(\sum_{n=1}^{\infty} q^n) = p\frac{d}{dq}(\frac{1}{1-q}) = p\frac{d}{dq}(1-q)^{-1} = p(1-q)^{-2} = \frac{1}{p} \end{aligned}$$

Now, we see a simple connection between $(S_n, n \in \mathbb{N})$ and $(T_k, k \in \mathbb{N})$

Note: For instance, $S_m \perp S_n - S_m \forall m < n$ i.e $(Z_1 + \cdots + Z_m) \perp (Z_{m+1} + \cdots + Z_n)$

Proposition 11.2. $T_n > n \iff X_n < k, \quad k \in \mathbb{N}, n \in \mathbb{N}$

then $T_k \sim negbin(k, p) \iff X_n \sim bin(n, p)$

Proof. ($\Longleftarrow$) we suppose $X_n \sim bin(n, p) \forall n$

$$\begin{aligned} P(T_k = n) &= P(T_k \leq n) - P(T_k < n) \\ &= P(T_k \leq n) - P(T_k \leq n-1) \\ &= P(T_k > n-1) - P(T_k > n) \\ &= P(X_n - 1 < k) - P(X_n < k) \\ &= P(S_n - 1 < k) - P(S_n < k) \\ &= P[(S_{n-1} < k)(S_n < k)^c] \\ &= P(S_{n-1} < k \leq S_n) \\ &= P(S_n < k-1, Z_n = 1) \end{aligned}$$

($\Longrightarrow$)

$$
\begin{aligned}
P(X_n = k) &= P(X_n \leq k) - P(X_n < k) \\
&= P(X_n < k+1) - P(X_n < k) \\
&= P(T_{k+1} > n) - P(T_k > n) \\
&= P(T_k \leq n) - P(Tk+1 \leq n) \\
&= P(T_k \leq n-1) - P(T_{k+1} \leq n-1) + P(T_k = n) - P(T_{k+1} = 0) \\
&= P(X_{n-1} = k) + P(T_k = n) - P(T_{k+1} = n) \\
&\overset{induction}{=} P(S_{n-1} = k) + P(T_k = n) - P(T_{k+1} = n) \\
&\overset{defn}{=} P(S_{n-1} = k) + P(S_{n-1} = k-1, Z_n = 1) - P(S_{n-1} = k, Z_n = 1) \\
&= P(S_{n-1} = k, Z_n = 1) + P(S_{n-1} = k, Z_n = 0) + P(S_{n-1} = k-1, Z_n = 1) - P(S_{n-1} = k, Z_n = 1) \\
&= P(S_n = k)
\end{aligned}
$$

□

Thus, $P(T_n > k) = P(S_n < k) = \sum_{i=0}^{k-1} \binom{n}{i} p^i q^{n-i}$

Note

$$
\begin{aligned}
W \sim geo(p) &\iff W \overset{d}{=} T \\
P(W = n) &= pq^{n-1} \quad n = 1, 2, cdots \\
EW &= \sum_{n=1}^{\infty} nP(W = n) = \sum_{n=1}^{\infty} npq^{n-1} \\
&= p\frac{d}{dq}(\sum_{n=0}^{\infty} q^n) = p\frac{d}{dq}(1-q)^{-1} = p(1-q)^{-2} = \frac{1}{p} \\
EW(W-1) &= EW^2 - EW \\
&= \sum_{n=1}^{\infty} n(n-1)P(W = n) = pq(\sum_{n=1}^{\infty} n(n-1)q^{n-2}) \\
&= pq + 2(1-q)^{-3} = \frac{2q}{p^2} \\
\text{Thus, } EW^2 &= \frac{2q}{p^2} + EW = \frac{p+2q}{p^2} = \frac{1-q}{p^2} \\
\text{Therefore, } Var(W) &= \frac{1-q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}
\end{aligned}
$$

12 Poisson Distribution

$$\binom{n}{k}p^k(1-p)^{n-k} = \frac{\frac{n(n-1)\cdots(n-k+1)}{n^k}}{(1-\frac{\lambda}{n})^k}\frac{\lambda^k}{k!}(1-\frac{\lambda}{n})^n$$
$$\lambda = np$$
$$p = \frac{\lambda}{n}$$
$$\frac{\frac{n(n-1)\cdots(n-k+1)}{n^k}}{(1-\frac{\lambda}{n})^k} = \frac{1(1-\frac{1}{n})\cdots(\frac{k-1}{n})}{(1-\frac{\lambda}{n})^k} \to 1, \text{Since } n \to \infty$$
$$(1-\frac{\lambda}{n})^n = e^{-\lambda}$$

Thus we define

Definition 12.1. $N \sim Poisson(\lambda) \iff P(N=k) = lim_{n\to\infty}P(X_n = k) = \frac{\lambda^k}{k!}e^{-\lambda}$ where $X_n \sim bin(n, \frac{\lambda}{n}), (N_t, t>0)$

12.1 N_t

N_t = random number of successes in time $t > 0$ and we assume:

- $N_t \sim Poisson(EN_t)$
- EN_t proportional to t

Note: $EN = \sum_{k=0}^{\infty} kP(N=k) = \sum_{k=1}^{\infty}(\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda})\lambda = \lambda$
$(N_t, t > 0)$ is a poisson distribution iff

- $N_t \sim Poisson(\lambda_t).t > 0$
- For any $0 < t_n \uparrow\uparrow$ as $n \to \infty$ (strictly increasing), $N_{t_1}, N_{t_2} - N_{t_1} \cdots N_{t_n} - N_{t_{n-1}} \cdots$ are mutually statistical independent.

12.2 T_n

$T_n(n \in \mathbb{N})$ = random amount of time for n successes

$$T_n > t \iff N_t < n$$

Thus,

$$1 - F_n(t) = P(T_n > t) = P(N_t < n)$$
$$= [\sum_{k=0}^{n-1}\frac{(\lambda t)^k}{k!}]e^{-\lambda t}$$

so $f_n(t) = F_n'(t) = \lambda \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}$

But let $Z_n = \lambda T_n$, to find the pdf for Z_n

$$\begin{aligned} g_n(Z) &= f_n(t)|\frac{dt}{dz}| = f_n(Z/\lambda)\frac{1}{\lambda} \\ &= \frac{Z^{n-1}e^{-Z}}{(n-1)!} = \frac{Z^{n-1}e^{-Z}}{\Gamma(n)} \end{aligned}$$

13 Gamma Distribution

Definition 13.1. $Z \sim Gamma(p), p > 0 \iff pdf : \frac{Z^{p-1}e^{-Z}}{\Gamma(p)}$ where $\Gamma(p) = \int_0^\infty Z^{p-1}e^{-Z}dz$

13.1 $\Gamma(p)$

- $\Gamma(p+1) = p\Gamma(p)$
- $\Gamma(n) = (n-1)!$
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

13.2 Gamma and Normal Distribution

$Z \sim N(0,1) \iff -Z \stackrel{d}{=} Z, \frac{Z^2}{2} \sim G(\frac{1}{2})$

Proof. ($\implies$)

$$\begin{aligned} let\ w &= Z^2/2 \\ G(W) &= P(W \le w) = P(-\sqrt{2w} \le Z \le \sqrt{2w}) \\ &= 2P(0 \le Z \le \sqrt{2w}) \\ &= 2[P(Z \le \sqrt{2w}) - P(Z \le 0)] \\ &= 2P(Z \le \sqrt{2w}) - 1 \\ G(W) &= 2\Phi(\sqrt{2w}) - 1 \\ \text{Thus, } g(w) &= G'(W) = 2\phi\sqrt{2w}\frac{\sqrt{2}}{2}w^{1/2-1} \\ &= \frac{1}{\sqrt{2\pi}}e^{-w}\sqrt{2}w^{1/2-1}, w > 0 \\ &= \frac{1}{\sqrt{\pi}}w^{1/2-1}, w > 0 \end{aligned}$$

□

Definition 13.2. $Z \sim G(p), p > 0 \iff g(Z) = \frac{Z^{p-1}e^{-z}}{\Gamma(p)}$

13.3 Expected Value and Variance

$$EZ^s = \frac{\Gamma(p+s)}{\Gamma(p)}$$

$Z \sim N(0,1)$

- n is odd $\implies -Z^n = (-Z)^n \overset{d}{=} Z^n \implies -EZ^n = EZ^n \implies EZ^n = 0$
- n is even $\implies$ since $Z^2 = 2W, W \sim G(\frac{1}{2}) \implies Z^n = Z^{2k} = 2^k W^k$

$$\begin{aligned} EZ^n &= 2^k EW^k \\ &= \frac{2^k \Gamma(k+\frac{1}{2})}{\Gamma(\frac{1}{2})} \\ &= 2^k \frac{\frac{2k-1}{2}\Gamma(\frac{2k-1}{2})}{\Gamma(\frac{1}{2})} \\ &= \text{Product of all the odd numbers below n} \end{aligned}$$

E.g $Z \sim N(0,1), EZ^6 = 15 \cdot 13 \cdot 11 \cdot 9 \cdot 7 \cdot 5 \cdot 3$ for n=odd

- $Var(Z^2) = EZ^4 - (EZ^2)^2 = 3 - 1 = 2$

Proposition 13.1. *$T = Z + W, U = Z/T$ then $Z \sim G(p), W \sim G(p), Z \perp W \iff$*

- $T \sim G(p+q)$
- $f_u(u) = \frac{\Gamma(p+q)}{\Gamma(p)+\Gamma(q)} u^{p-1}(1-u)^{q-1}, 0 < u < 1$
- $T \perp U$

Proof.

$$\begin{aligned} g(z,w) &= \frac{Z^{p-1}e^{-Z}W^{q-1}e^{-W}}{\Gamma(p)\Gamma(q)} \\ h(u,t) &= g(z,w)|\frac{\partial(z,w)}{\partial(u,t)}|_+ \\ |\frac{\partial ut, (1-u)t}{\partial u, t}|_+ &= \begin{bmatrix} t & u \\ -t & 1-u \end{bmatrix}_+ = t \\ h(u,t) &= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} u^{p-1}(1-u)^{q-1} \frac{t^{p+1-1}e^{-t}}{\Gamma(p+q)} \end{aligned}$$

□

14 Beta Distribution

Definition 14.1. $U \sim beta(p,q) \iff U = \frac{Z}{Z+W}, Z \sim G(p), W \sim G(P), Z \perp W$

Notice that with

$$
\begin{aligned}
T = Z + W &\implies Z^2 = U^2 T^2 \\
\text{so, } EZ^2 &= EU^2 ET^2, \text{using independent} \\
\text{so, } EU^2 &= \frac{EZ^2}{ET^2} = \frac{\Gamma(p+2)/\Gamma(p)}{\Gamma(p+q+2)/\Gamma(p+q)} \\
EU &= \frac{EZ}{ET} = \frac{p}{p+q} \\
Z &\sim G(p), EZ^s = \Gamma(p+s)/\Gamma(p) \\
EZ &= p = Var(Z) \\
EZ^{-1} &= \Gamma(p-1)/\Gamma(p) = \frac{1}{p-1} \\
E(Z/W) &= E(ZW^{-1}) = EZEW^{-1} = p\frac{1}{q-1} = \frac{p}{q-1}
\end{aligned}
$$

15 χ^2 Distribution

$\chi^2_{(m)} = 2Z, Z \sim G(m/2)$

Since $Z_1, \cdots, Z_n$ IID $Z \sim N(0,1)$ and $\frac{Z_1^2}{2}, \cdots, \frac{Z_n^2}{2}$ IID $\frac{Z^2}{2} \sim G(\frac{1}{2}) \implies \frac{2\sum Z_i^2}{2} \sim 2G(\frac{n}{2}) = \chi^2_{(n)}$

16 Additional Definitions

Definition 16.1. $Z \sim N(0,1), \phi(Z) = \frac{1}{\sqrt{2\pi}} e^{-Z^2/2}$

Definition 16.2. $X \sim N(\mu, \sigma^2) \iff X = \mu + \sigma Z$

Definition 16.3. $Z \sim exp(1) \iff f(Z) = e^{-Z}$

Definition 16.4. $X \sim exp(\theta) \iff X = \theta Z$

Definition 16.5. $Z \sim G(p) \iff f(Z) = \frac{z^{p-1}e^{-Z}}{\Gamma(p)}$ and $X \sim G(p,\theta) \iff X = \theta Z$

Thus, $EX^s = E(\theta Z)^s = \theta^s EZ^s = \frac{\theta^s \Gamma(p+s)}{\Gamma(p)}$