Math 170B: Probability Theory II

Prof. R. Killip | Spring 2024

Contents

1 Characteristic & Moment-Generating Functions

Definition. Given a random variable X, the ***characteristic function*** [CF] of X is the function $\phi_X : \mathbb{R} \to \mathbb{C}$ defined by:

$$\boxed{\phi_X(\xi) = \mathbb{E}\left(e^{i\xi X}\right)}$$

Definition. Given a random variable X, the ***moment-generating function*** [MGF] of X is the function $M_X : \mathbb{R} \to \mathbb{R}$ defined by:

$$\boxed{M_X(t) = \mathbb{E}\left(e^{tX}\right)}$$

Characteristic & Moment-Generating Functions

Suppose $M_X(t)$ exists in some neighborhood $[-\delta, \delta]$ of 0 ($\delta > 0$); then $\forall\, t \in [-\delta, \delta]$:

1. $M_X(t)$ can be expressed as a power series:

$$\boxed{M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbb{E}\left(X^k\right)}$$

2. Can derive the moments of X from the characteristic/moment-generating functions:

$$\boxed{\mathbb{E}\left(X^k\right) = \left.\frac{d^k}{dt^k} M_X(t)\right|_{t=0} = \left.(i)^{-k} \frac{d^k}{d\xi^k}\phi_X(\xi)\right|_{t=0}}$$

3. For $\xi \in [-\delta, \delta]$ ($|\xi| \leq \delta$), can treat characteristic function, MGF similarly:

$$\boxed{\phi_X(\xi) = M_X(i\xi)}$$

4. ϕ_X is analytic: for every point $\xi \in \mathbb{R}$, ϕ_X has value given in some neighborhood of ξ by power series with radius of convergence $r \geq \delta$

Existence:
- Characteristic function $\phi_X(\xi)$ defined/exists $\forall\, \xi \in \mathbb{R}$
  - In particular: if $\mathbb{E}\left(|X|^n\right) < \infty$, then $\phi_X(\xi)$ is n-times differentiable at the origin and $\phi_X^{(n)}(0) = i^n \mathbb{E}\left(X^n\right)$.
- Moment-generating function $M_X(t)$ strictly positive, only guaranteed to exist at $t = 0$; may not be defined for $t \in \mathbb{R} \setminus \{0\}$
  - k^{th} moment of X exists iff $M_X(t)$ is k-times differentiable at $t = 0$

1.1 Multivariate Characteristic Functions

Definition. Let $X_1, X_2, \ldots, X_n$ be random variables; then the ***joint characteristic function*** of $\vec{X} = \langle X_1, X_2, \ldots, X_n \rangle$ is given by:

$$\boxed{\phi_{\vec{X}}\left(\vec{\xi}\right) = \mathbb{E}\left(e^{i\xi_1 X_1 + i\xi_2 X_2 + \ldots + i\xi_n X_n}\right)}$$

Theorem (*Joint Characteristic Functions*).

$$X_1, X_2, \ldots, X_n \text{ independent} \iff \phi_{\vec{X}}\left(\vec{\xi}\right) = \phi_{X_1}(\xi_1) \cdot \phi_{X_2}(\xi_2) \cdot \ldots \cdot \phi_{X_n}(\xi_n)$$

Multivariate Characteristic Functions

(i) The joint characteristic function $\phi_{\vec{X}}\left(\vec{\xi}\right)$ uniquely determines the joint law of $\vec{X}$

(ii) If $X_1, X_2, \ldots, X_n$ are independent, then the characteristic function of $Y = X_1 + X_2 + \ldots + X_n$ is:

$$\boxed{\phi_Y(\xi) = \prod_{i=1}^{n} \phi_{X_i}(\xi)}$$

1.2 Cumulants

Definition. Given a random variable X, the ***cumulant-generating function*** [CGF] of X is defined (for $|\xi|$ small) as:

$$\boxed{K : \xi \mapsto \ln[\mathbb{E}\left(e^{i\xi X}\right)] = \ln[\phi_X(\xi)]}$$

Cumulants:

1. $\kappa_1 : \kappa_1 = \frac{1}{i}\frac{d}{d\xi}\ln[\phi_X(\xi)]\Big|_{\xi=0} = \mathbb{E}(X)$
2. $\kappa_2 : \kappa_2 = \frac{1}{i^2}\frac{d^2}{d\xi^2}\ln[\phi_X(\xi)]\Big|_{\xi=0} = \mathrm{var}(X)$

1.3 Bochner's Theorem

Recall. A symmetric/Hermitian matrix $A \in M_{n\times n}(\mathbb{C})$ is called:

1. ***Positive semi-definite*** if $z^T A z \geq 0 \; \forall \; z \in \mathbb{C}^n$
2. ***Positive definite*** if $z^T A z > 0$ strictly.

<u>**Properties**</u>:

(i) A is positive semi-definite iff $\lambda \geq 0$ real non-negative for every eigenvalue λ of A. In particular, A is positive definite iff $\lambda > 0$.

- Corollary: $\det(A) \geq 0$ [semi-definite] / $\det(A) > 0$, A invertible [definite]

(ii) A is positive semi-definite iff $\exists \; L \in M_{n\times n}(\mathbb{C})$ such that $A = L^* L$. In particular, A is positive definite iff L is invertible.

Theorem (*Bochner's Theorem*). A function $\phi : \mathbb{R} \to \mathbb{C}$ is a characteristic function for some r.v. X iff:

1. $\phi(0) = 1$
2. ϕ is continuous
3. ϕ is "positive-definite": $\sum_{i,j=1}^{n} \overline{\lambda}_i \lambda_j \phi(\xi_i - \xi_j) \geq 0 \; \forall \; n \in \mathbb{N}, \xi_i \in \mathbb{R}, \lambda_i \in \mathbb{C}$

   Equivalently, the matrix $\big(\phi(\xi_i - \xi_j)\big)_{ij}$ is positive-definite, where:

$$\big(\phi(\xi_i - \xi_j)\big)_{ij} = \begin{pmatrix} \phi(0) & \phi(\xi_2 - \xi_1) & \dots & \phi(\xi_n - \xi_1) \\ \phi(\xi_1 - \xi_2) & \phi(0) & \dots & \phi(\xi_n - \xi_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\xi_1 - \xi_n) & \phi(\xi_2 - \xi_n) & \dots & \phi(0) \end{pmatrix}$$

2 Applications of Characteristic Functions

2.1 Uniqueness Theorem

Theorem (***Characterizations of a Distribution***)**.** Let X, Y be two random variables; then TFAE:

1. $F_X(x) = F_Y(x) \ \forall \ x \in \mathbb{R}$
2. $\phi_X(\xi) = \phi_Y(\xi) \ \forall \ \xi \in \mathbb{R}$
3. $\mathbb{E}(f(X)) = \mathbb{E}(f(Y))$ for all bounded continuous functions $f : \mathbb{R} \to \mathbb{R}/\mathbb{C} \to \mathbb{C}$

2.2 Moment Problem

Moment Problem:

- If an r.v. X has any moment $\mathbb{E}\left(X^k\right) = \infty$, then its distribution F_X is not uniquely determined by its moments.
- If F_X is determined by finitely many moments, then X takes only finitely many values.

(∗) All moments $\mathbb{E}\left(X^k\right)$ of X exist & are finite iff all even moments exist & are finite

Theorem (***Uniqueness Theorem for Moments***)**.** Let $m_k = \mathbb{E}\left(X^k\right)$; then the moments m_k characterize X if, for some $a > 0$:

$$\boxed{\sum_{k=0}^{\infty} \frac{a^{2k}}{(2k)!} m_{2k} < \infty}$$

3 Branching Processes

Definition. Given a random variable X, the ***probability-generating function*** of X is the function $G_X : \mathbb{R} \to \mathbb{R}$ defined by:

$$\boxed{G_X(z) = \mathbb{E}\left(z^X\right)}$$

Prop. Let N be an r.v. with values in $\mathbb{N} \cup \{0\}$ and let $X_1, X_2, \ldots, X_j, \ldots$ i.i.d. and independent of N. Let $Y = X_1 + X_2 + \ldots + X_N$; then:

$$\boxed{\phi_Y(\xi) = \phi_N\left(\frac{1}{i}\ln[\phi_X(\xi)]\right)}$$

Corollary:

(a) If X and N admit MGFs, have $M_Y(t) = M_N(\ln[M_X(t)])$

(b) $\mathbb{E}(Y) = \mathbb{E}(X) \cdot \mathbb{E}(N)$

(c) $\operatorname{var}(Y) = \operatorname{var}(N)\,\mathbb{E}(X)^2 + \mathbb{E}(N)\operatorname{var}(X)$

(d) $G_Y(z) = G_n \circ G_X(z)$

Theorem (*Galton-Watson Tree*). Suppose an individual [gen 0] has X-many children, each of whom has independently and with distribution X and so on ad infinitum.

Let Z_n denote the number of children in the n^{th} generation [$Z_0 = 1$]; then:

1. $\mathbb{E}(Z_n) = \mathbb{E}(X)^n$
2. Let $e_n = \mathbb{P}(Z_n = 0)$; then $e_{n+1} \geq e_n$. In particular, $e = \lim_{n\to\infty} e_n$ exists.
3. $e \in [0,1]$ is the smallest root/solution to $G_X(x) = x$ for $x \in [0,1]$.
4. $e < 1$ iff either (i) $\mathbb{E}(X) > 1$ or (ii) $X \equiv 1$.
5. $e = 0$ iff $\mathbb{P}(X = 0) = 0$.

4 Conditional Expectation

Theorem (***Conditional Expectation***)**.** Let X, Y random variables and $\mathbb{E}(|Y|) < \infty$; then there is a unique r.v. $\mathbb{E}(Y \mid X)$ [finite] satisfying:

(i) $\mathbb{E}(Y \mid X)$ depends only on X, i.e. $\mathbb{E}(Y \mid X) = g(X)$ for some function g

(ii) For any bounded function h, have that:

$$\boxed{\mathbb{E}(Y \cdot h(X)) = \mathbb{E}(\mathbb{E}(Y \mid X) \cdot h(X))}$$

Recall. $\mathbb{E}(Y \mid X) = g(X)$ defined via:

$$\mathbb{E}(Y \mid X) = \sum_{n \in \mathbb{Z}} 1_{\{X=n\}}(\omega)\mathbb{E}(Y \mid X = n) \qquad [X \text{ discrete}]$$

$$\mathbb{E}(Y \mid X) = \frac{\int y f_{X,Y}(x, y)\, dx}{\int f_{X,Y}(x, y)\, dy} \qquad [X \text{ continuous}]$$

Theorem (***Conditional Expectation***)**.** Let X, Y r.v.s with $\mathbb{E}(|Y|) < \infty$:

(i) **Law of Iterated Expectation**:

$$\boxed{\mathbb{E}(\mathbb{E}(Y \mid X)) = \mathbb{E}(Y)}$$

(ii) Define the ***conditional variance*** of Y given X:

$$\boxed{\operatorname{var}(Y \mid X) := \mathbb{E}\left([Y - \mathbb{E}(Y \mid X)]^2 \mid X\right) = \mathbb{E}\left(Y^2 \mid X\right) - \mathbb{E}(Y \mid X)^2}$$

(iii) Can obtain the **Law of Total Variance**:

$$\boxed{\operatorname{var}(Y) = \mathbb{E}(\operatorname{var}(Y \mid X)) + \operatorname{var}(\mathbb{E}(Y \mid X))}$$

5 Inequalities

Theorem (***Markov's Inequality***)**.** Let X non-negative; then for any $\lambda > 0$

$$\boxed{\mathbb{P}(X \geq \lambda) \leq \frac{1}{\lambda}\mathbb{E}(X)}$$

Corollary (***Chebyshev's Inequality***)**.** Let X r.v. and $\mu \in \mathbb{R}$; then:

$$\boxed{\mathbb{P}(|X - \mu| \geq \lambda) \leq \frac{1}{\lambda^2}\mathbb{E}\left((X - \mu)^2\right)}$$

In particular, if $\mu = \mathbb{E}(X)$:

$$\boxed{\mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda) \leq \frac{1}{\lambda^2}\mathrm{var}(X)}$$

Corollary (***Exponential Markov***)**.** For any $s \geq 0$, have:

$$\mathbb{P}(X \geq \lambda) \leq e^{-s\lambda}\mathbb{E}\left(e^{sX}\right)$$

Large Deviation Argument (***Cramer's Method***): To obtain a smaller upper bound for $\mathbb{P}(X \geq \lambda)$, can solve for $e^{-s\lambda}\mathbb{E}\left(e^{sX}\right)$ and minimize over all s.

Theorem (***Cauchy-Schwarz Inequality***)**.**

$$\boxed{\mathbb{E}(XY)^2 \leq \mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right)}$$

Theorem (***Jensen's Inequality***)**.** For any convex function ϕ:

$$\boxed{\phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X))}$$

6 Modes of Convergence

Definition. Say that random variables X_n ***converge in distribution*** to a random variable X if, $\forall$ bounded continuous functions f:

$$\boxed{\mathbb{E}\left(f(X_n)\right) \xrightarrow{n\to\infty} \mathbb{E}\left(f(X)\right)}$$

Theorem (***Levy-Cramer***)**.** Given random variables X_n and X, TFAE:

1. $\mathbb{E}\left(f(X_n)\right) \to \mathbb{E}\left(f(X)\right)$ $\forall$ bounded & continuous f
2. $F_{X_n}(x) \to F_X(x)$ for every $x \in \mathbb{R}$ at which $F_X(x)$ is continuous
3. $\phi_{X_n}(\xi) \to \phi_X(\xi)$ $\forall\, \xi \in \mathbb{R}$

Mode of Convergence

Say that a sequence of r.v.s X_n converge to an r.v. X [on the same probability space]:

1. ...**in p-mean/L$^\mathbf{p}$** $(1 \leq p < \infty)$ if:

$$\boxed{\mathbb{E}\left(|X - X_n|^p\right) \to 0 \text{ as } n \to \infty}$$

2. ...**in probability** if, $\forall\, \epsilon > 0$:

$$\boxed{\mathbb{P}\left(|X_n - X| > \epsilon\right) \to 0 \text{ as } n \to \infty}$$

3. ...**almost surely** [a.s.] if:

$$\boxed{\mathbb{P}\left(\{\omega : X_n(\omega) \to X(\omega\}\right) = 1}$$

Theorem. Fix $1 \leq p \leq r < \infty$. Given r.v.s X_n, $X : \Omega \to \mathbb{R}$, the modes of convergence of X_n to X are interrelated as follows:

$$r\text{-mean convergence} \implies p\text{-mean convergence} \implies \text{convergence in } \mathbb{P} \tag{1}$$
$$\text{convergence almost surely} \implies \text{convergence in } \mathbb{P} \tag{2}$$
$$\text{convergence in } \mathbb{P} \implies \text{convergence in distribution} \tag{3}$$

6.1 Compactness for Convergence in Distribution

Definition. A family of r.v.s X_n is called ***tight*** if, for every $\epsilon > 0$, $\exists\, L > 0$ s.t.:

$$\boxed{\mathbb{P}\left(|X_n| \geq L\right) < \epsilon \;\forall\, n \in \mathbb{N}}$$

Theorem (***Prokhorov***)**.** If a sequence of r.v.s X_n are tight, then there is a subsequence X_{n_k} that converges in distribution.

Theorem (***Helly selection theorem***)**.** Given a sequence F_n of non-decreasing cadlag functions that are uniformly bounded on a bounded interval $[a, b)$, there is:

1. A subsequence F_{n_k} of F_n, and
2. A non-decreasing cadlag function $F : [a, b) \to \mathbb{R}$

such that $F_{n_k} \to F$ pointwise [for points where F is continuous].

Proof. Can enumerate the rationals by q_m; at each q_m, find a subsequence F_{n_k} s.t. $F_{n_k}(q_m)$ converges. Via the Cantor diagonal slash argument, find the subsequence s.t. $F_{n_k}(q_m)$ converges $\forall\, q_m$ and take F to be the limit of F_{n_k}. □

Lemma (***Borel-Cantelli***)**.** If the sum of probabilities of events E_n is finite [i.e. $\sum_{n=1}^{\infty} \mathbb{P}\left(E_n\right) < \infty$], then the probability of infinitely many E_n's occurring is 0:

$$\mathbb{P}\left(\limsup_{n\to\infty} E_n\right) = 0$$

7 The Central Limit Theorem

Theorem (***Central Limit Theorem***)**.** Let X_i be i.i.d. random variables with mean $\mathbb{E}\left(X_i\right) = \mu$ and finite variance $\operatorname{var}\left(X_i\right) = \sigma^2 < \infty$. Then:

$$\boxed{Z_n = \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{\text{in dist.}} \mathrm{N}\left(0,\, 1\right)}$$

Proof. Via characteristic functions:

$$\phi_{Z_n}\left(\xi\right) \to \phi_{\mathrm{N}(0,\,1)}\left(\xi\right) \text{ pointwise } \forall\, \xi \in \mathbb{R}.$$

$\square$

Theorem (***Berry-Esseen Theorem***)**.** Suppose X_i are i.i.d. r.v.s with $\mathbb{E}\left(|X_i|^3\right) < \infty$, and define the Z_n's by

$$Z_n = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Then the CDFs F_n of the Z_n's satisfy that:

$$\boxed{\sup_z |F_n(z) - \Phi(z)| \leq \frac{1}{\sqrt{n}} \frac{\mathbb{E}\left(|X_i|^3\right)}{\sigma^3}}$$

8 Laws of Large Numbers

8.1 The Weak Law of Large Numbers

Theorem. Suppose that $\operatorname{var}(X_i) < \infty$; then:

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow[n \to \infty]{L^2} \mathbb{E}(X_i)$$

Lemma. If a sequence of random variables Z_n converge $Z_n \xrightarrow{d} C$ to some $C \in \mathbb{R}$, then $Z_n \to C$ in probability as well.

Theorem (***Weak/Khinchin LLN***)**.** If X_i are i.i.d. r.v.s with $\mathbb{E}(|X_i|) < \infty$, then:

$$\boxed{\frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow[n \to \infty]{p} \mathbb{E}(X_i)}$$

Proof. Can show that:

$$Z_n = \frac{X_1 + \ldots + X_n - n\mu}{n} \implies \phi_{Z_n}(\xi) \to \phi_0(\xi) \text{ pointwise } \forall\, \xi \in \mathbb{R}$$

and apply Lemma. □

8.2 Strong Laws of Large Numbers

Theorem (***Kolmogorov Maximal Theorem***)**.** Let W_i i.i.d. r.v.s with $\mathbb{E}(W_i) = 0$, $\mathbb{E}(W_i^2) < \infty$. Then for each $N \in \mathbb{N}$, $\lambda > 0$:

$$\boxed{\mathbb{P}\left(\sup_{1 \leq n \leq N} \left| \sum_{i=1}^{n} W_i \right| > \lambda\right) \leq \frac{\sum_{i=1}^{N} \operatorname{var}(W_i)}{\lambda^2}}$$

In particular, sending $N \to \infty$:

$$\boxed{\mathbb{P}\left(\sup_{n \in \mathbb{N}} \left| \sum_{i=1}^{n} W_i \right| > \lambda\right) \leq \frac{1}{\lambda^2}\left(\sum_{i=1}^{\infty} \operatorname{var}(W_i)\right)}$$

Lemma (**Kronecker**)**.** Let $y_k \in \mathbb{R}$ such that $\sum_{k=1}^{\infty} \frac{y_k}{k}$ converges. Then:

$$\boxed{\frac{y_1 + y_2 + \ldots + y_n}{n} \to 0}$$

Theorem (***Strong/Kolmogorov LLN***)**.** If X_i are i.i.d. random variables with $\mathbb{E}(|X_i|) < \infty$, then:

$$\boxed{\frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow[n \to \infty]{\text{almost surely}} \mathbb{E}(X_i)}$$

Proof. (*Finite 4^{th} Moment*) Via Chebyshev's inequality.

(*Finite Variance*) Via Kronecker's Lemma.

(*Kolmogorov LLN*) Define $X_n = X_n' + X_n''$, $X_n' = 1_{|X_n| \leq n} X_n$, $X_n'' = 1_{|X_n| > n} X_n$.

Can split up original expression:

$$\frac{X_1 + \ldots + X_n}{n} = \frac{[X_1' - \mathbb{E}(X_1')] + \ldots + [X_n' - \mathbb{E}(X_n')]}{n} + \frac{\mathbb{E}(X_1') + \ldots + \mathbb{E}(X_n')}{n} + \frac{1}{n}\sum_{k=1}^{n} X_k''$$

and show separately that all three terms $\to 0$ almost surely. $\square$

9 Random Walks

9.1 Simple Random Walks

Setup: Let X_k be i.i.d. random variables defined by:

$$\boxed{X_k = \begin{cases} +1 & \text{with probability } p \\ -1 & \text{with probability } q = 1-p \end{cases}}$$

and define for $n \in \mathbb{N}$:

$$\boxed{S_n = \sum_{k=1}^{n} X_k}$$

Lemma. Define:

$$\boxed{u_n = \mathbb{P}\left(\text{return to origin at time } n\right) = \mathbb{P}\left(S_n = 0\right)}$$

Then $\forall\ m \in \mathbb{N}$:

$$u_{2m+1} = 0; \quad u_{2m} = \binom{2m}{m} p^m (1-p)^m$$

Proof. Trivial. □

Lemma. For $z \in \mathbb{R}$ with $|z|^2 \leq \frac{1}{4pq}$:

$$\boxed{U(z) := \sum_{n=0}^{\infty} u_n z^n = \frac{1}{\sqrt{1-4pqz^2}}}$$

Proof. Can keep taking derivatives $\frac{d}{d\epsilon}(1-\epsilon)^{-(m-\frac{1}{2})}$ to find that

$$(1-\epsilon)^{-\frac{1}{2}} = 1 + \sum_{m=1}^{\infty} \frac{(2m-1)\cdot \ldots \cdot 3 \cdot 1}{2^m m!} \epsilon^m$$

and substitute $\epsilon = 4pqz^2$. □

Theorem. Define:

$$\boxed{f_n = \mathbb{P}\left(\text{first return to origin at time } n\right) = \mathbb{P}\left(S_n = 0 \text{ but } S_{n-1}, \ldots, S_2, S_1 \neq 0\right)}$$

Then:

$$\boxed{F(z) := \sum_{n=0}^{\infty} f_n z^n = 1 - \frac{1}{U(z)}}$$

In particular:

$$f_{2m+1} = 0; \quad f_{2m} = \frac{1}{2m-1}\binom{2m}{m} p^m (1-p)^m$$

Proof. Can use that:

$$u_n = f_n + \sum_{m=1}^{n-1} f_m u_{n-m}$$

and substitute into expression for $U(z)$. $\square$

Corollary.

$$\boxed{\mathbb{P}\left(\text{eventually return to origin}\right) = 1 - |p-q|}$$

Prop. Let T_1 = time of 1^{st} return; $T_1 = \min\{n : S_n = 0\}$ for a symmetric SRW $[p = q]$. Then $\mathbb{E}\left(T_1\right) = \infty$.

9.2 The Gambler's Ruin Problem

Setup: Assign values $V(0) = 0$, $V(N) = 1$; a person starting at $0 < n < N$ performs a simple random walk ($\mathbb{P}(\text{step right}) = p$, $\mathbb{P}(\text{step left}) = q$) until reaching either 0 ("lose") or N ("win").

$\Rightarrow$ Want to find the "value" $V(n) = \mathbb{E}(V \,|\, \text{start at } n)$ of a position $0 < n < N$.

1. Boundary conditions: $V(0) = 0, V(N) = 1$
2. Difference equation: $v(n) = pv(n+1) + qv(n-1)$

Symmetric SRW (*simple case*: $p = q = 1/2$):

$$\boxed{v(n) = \frac{v(n+1) + v(n-1)}{2} \implies v(n) = \frac{n}{N}}$$

Biased SRW (*hard case*: $p \neq q$):

$$\boxed{v(n) = \frac{\left(\frac{q}{p}\right)^n - 1}{\left(\frac{q}{p}\right)^N - 1}}$$

Can compute *expected game length* $e(n)$:

$$\boxed{e(n) = \begin{cases} n \cdot (N - n) & p = q = 1/2 \\ \frac{1}{q-p}\left[N\frac{\left(\frac{q}{p}\right)^n - 1}{\left(\frac{q}{p}\right)^N - 1} - n\right] & p \neq q \end{cases}}$$

10 Appendix

10.1 Distributions of Discrete Random Variables

Name	Parameters	Distribution	Characteristic
Bernoulli(p)	$p \in [0,1]$	$f_X(k) = \begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$	$1-p+pe^{i\xi}$
Binomial(n,p)	$n \in \mathbb{N}, p \in [0,1]$	$f_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & k=0,1,\ldots,n \\ 0 & \text{otherwise} \end{cases}$	$(1-p+pe^{i\xi})^n$
Geometric(p)	$p \in (0,1]$	$f_X(k) = \begin{cases} (1-p)^{k-1}p & k=1,2,3,\ldots \\ 0 & \text{otherwise} \end{cases}$	$\frac{pe^{i\xi}}{1-(1-p)e^{i\xi}}$
Poisson(λ)	$\lambda \geq 0$	$f_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$	$e^{\lambda(e^{i\xi}-1)}$

10.2 Distributions of Continuous Random Variables

Definition (***Gamma Function***)**.** The ***gamma function*** is the function defined by:

$$\boxed{\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt}$$

Name	**Parameters**	**Distribution**	**Characteristic**
$\mathrm{N}\left(\mu,\, \sigma^2\right)$	$\mu, \sigma \in \mathbb{R}$	$f_X\left(x\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	$e^{i\mu\xi - \frac{1}{2}\sigma^2\xi^2}$
$\mathrm{N}\left(\vec{\mu},\, \Sigma\right)$	$\vec{\mu} \in \mathbb{R}^k, \Sigma \in \mathbb{R}^{k\times k}$	$\frac{1}{(2\pi)^{k/2}\det(\Sigma)^{\frac{1}{2}}} \exp\left(-\frac{(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})}{2}\right)$	$e^{i\vec{\xi}\cdot\vec{\mu} - \frac{1}{2}\vec{\xi}\cdot\Sigma\vec{\xi}}$
$\Gamma(k, \lambda)$	$k, \lambda > 0$	$f_X\left(x\right) = \begin{cases} \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$	$(1 - \frac{i\xi}{\lambda})^{-k}$
$\mathrm{Exp}(\lambda)$	$\lambda > 0$	$f_X\left(x\right) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\frac{\lambda}{\lambda - i\xi}$