Final Review

Andrew Wang

August 13, 2024

1 Counting

Theorem. (Binomial Theorem) Let $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$. Then

$$(x+y)^n = \sum_{0 \leq i \leq n} \binom{n}{i} x^i y^{n-i}.$$

Theorem. (Multinomial Theorem) Let $a_1, \dots, a_n \in \mathbb{R}$ and $n \in \mathbb{N}$. Then

$$(a_1 + \cdots + a_k)^n = \sum_{m_1 + \cdots + m_k = n} \binom{n}{m_1, \dots m_k} a_1^{m_1} \dots a_n^{m_k}.$$

2 Probabilities

Definition. The *conditional probability* of E given F is

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$$

Definition. Two events E and F are said to be *independent* if

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F).$$

Theorem. (Bayes' Formula)

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(F|E)\mathbb{P}(E)}{\mathbb{P}(F)}.$$

3 Random Variables

Definition. The *variance* of a real valued random variable X is given by

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Definition. Two random variables X, Y are said to be independent if

$$\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b)$$

for all a, b.

Proposition. Let X, Y be two independent random variables. then:
(a) $g(X), f(Y)$ are independent for all functions g, f on X, Y.
(b) $\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[f(Y).]$
(c) $Var(g(X) + f(Y)) = Var(g(X)) + Var(f(Y))$

4 Poisson Random Variables

Definition. X is a *binomial random variable* X with parameter $p \in (0, 1)$ and $n \in \mathbb{N}$ is given by

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

for some $k \in \{1, 2, \ldots, n\}$. **Proposition.** Suppose $X_1, X_2, \ldots, X_N$ are independent Bernoulli random variables with parameter p. Then $X_1 + X_2 + \ldots X_n$ is a binomial random variable with parameters p, n.

Definition. X is a *Poisson random variable* with parameter $\lambda > 0$ if

$$\mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

for some integer k.

Proposition. Suppose $X = Poiss(\lambda_1), Y = Poiss(\lambda_2)$, then $X + Y = Poiss(\lambda_1 + \lambda_2) = Poiss(\lambda_1 + \lambda_2)$.

5 Continuous Random Variables

Definition. For $\mu \in \mathbb{R}, \sigma > 0$, X is a *normal random variable with mean* μ *and variance* σ^2 denoted by (μ, σ^2) if its probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

for some $x \in \mathbb{R}$.

6 Joint Distributions

Definition. For real valued random variables X, Y, the *joint cumulative distribution function* is $F : \mathbb{R} \times \mathbb{R} \to [0, 1]$ given by

$$F(a, b) = \mathbb{P}(X \leq a, Y \leq b).$$

with F being non-decreasing in each of its dimensions.

Definition. X,Y are *jointly continuous random variables* if there exists $f : \mathbb{R} \times \mathbb{R} \to [0, \infty]$ with $\int_{\mathbb{R}\times\mathbb{R}} f(x,y)dxdy = 1$

$$\mathbb{P}(X \in I_1, Y \in I_2) = \int_{I_1 \times I_2} f(x,y)dxdy$$

for any intervals I_1, I_2.

Proposition. Suppose X, Y are independent random variables, then
(a) $F(a,b) = F_X(a)F_Y(b)$,
(b) If X, Y are continuous with probability density function f_X, f_Y, respectively, then they are jointly continuous with joint probability distribution function $f(a,b) = f_X(a)f_Y(b)$. **Proposition.** Suppose X, Y are independent continuous random variables with probability density functions f_X, f_Y. Then $Z = X + Y$ is a continuous random variable with probability density function

$$g(a) = \int_{\mathbb{R}} f_X(x)f_Y(a-x)dx$$

Proposition. Suppose $X = \mathcal{N}(\mu_1, \sigma_1^2), Y = \mathcal{N}(\mu_2, \sigma_2^2)$ are independent, Then for $a, b \in \mathbb{R}$, $a^2 + b^2 > 0$,

$$aX + bY = \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

7 Moment Generation Functions

Definition. For a real-valued random variable X, its *moment-generating function* is $M_X : \mathbb{R} \to [0, \infty]$,

$$M_X(t) = \mathbb{E}[e^{tX}], t \in \mathbb{R}$$

Proposition.
(a) $M_X(0) = 1$.
(b) $M_{X+Y} = M_X M_Y$ for X, Y independent,
(c) M_X is fully determined by the moments of X so long as $\mathbb{E}[X^n] \leq (cn)^n$ for some $c > 0$.

Definition. For random variables $X_1, \ldots, X_n$, they have moment generating function $M_{X_{1,\ldots,n}}(t_1, \ldots, t_n) = \mathbb{E}[e^{t_1X_1+\cdots+t_nX_n}]$.

Definition. $X_1, X_2, \ldots, X_n$ are *multivariate normal random variable* if for any $X_i \in X_1, \ldots, X_n$ there exist independent normal random variables $Y_1, Y_2, \ldots, Y_m$ and $a_{ij} \in \mathbb{R}$ such that

$$\sum_{0 \leq j \leq m} a_{ij}Y_j = X_i.$$

8 Conditional Expectations and Inequaliies.

Definition. For a set S and $A \subset S$, the *indicator/characteristic function* of A is

$$\mathcal{X}_A(x) = \begin{cases} 1, x \in A, \\ 0, x \notin .A \end{cases}$$

Proposition. (Chevyshev's Inequality) For $a > 0$ and a real-valued random variable X,

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{Var(X)}{a^2}.$$

Proposition. (Markov's Inequality) For $a > 0$, and a real-valued random variable $X \geq 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

Definition. The *conditional expectation* of X given Y is a ramdom variable T with

$$\mathbb{P}(T = s) = \sum_{y \in Y : E[X|Y=y]=s} \mathbb{P}(Y = y)$$

and denoted by $\mathbb{E}[X|Y]$. Similarly, when X, Y are jointly continuous,

$$\mathbb{E}[X|Y = y] = \frac{1}{f_Y(y)} \int_{\mathbb{R}} x f(x, y) dx$$

$$\mathbb{P}(T \in I) = \int_{S(I)} f(y) dy$$

where $S(I) = y \in \mathbb{R} : \mathbb{E}[X|Y = y] \in I$

Proposition.
(a) $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$,
(b) $\mathbb{E}[Xg(Y)|Y] = g(Y)\mathbb{E}[X|Y]$.

9 Law of Large Numbers

Definition. X_n *converges in probability* to X if for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0$$

denoted by $X_n \xrightarrow{p} X$.

Definition. X_n *converges almost surely* to X if

$$\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1.$$

denoted by $X_n \xrightarrow{a.s.} X$.

Theorem. (Weak and Strong Law of Large Numbers) Suppose $\{X_n\}$ are i.i.d random variables with $X_n \xrightarrow{d} X, \mathbb{E}[X] < \infty$, then

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{p \ \& \ a.s.} \mathbb{E}[X].$$

10 Central Limit Theorem

Definition. $X_n \in \mathbb{R}$ *converges in distribution* to $X \in \mathbb{R}$ if

$$\lim_{n\to\infty} \mathbb{P}(X_n \leq a) = \mathbb{P}(X \leq a)$$

for all $a \in \mathbb{R}$ at which the CDF of X is continuous. This is denoted as

$$X_n \xrightarrow{d} X.$$

Theorem. (Central Limit Theorem) Suppose $X_n \in \mathbb{R}$ are i.i.d. with $\mathbb{E}[X^2] < \infty$, $Var(X) > 0$. Then

$$\frac{X_1 + \cdots + X_n - n\mathbb{E}[X]}{\sqrt{n \cdot Var(X)}} \Rightarrow \mathcal{N}(0,1).$$