1 Variance

Def

$Var(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$

Properties

1. $Var(X) \geq 0$
2. $Var(X) = 0 \iff P(X = E[X]) = 1$
3. $Var(X + c) = Var(X)$
4. $Var(aX) = a^2 Var(X)$
5. $Var(X + Y) = Var(X) + Var(Y)$
   (X, Y are independent)

2 PDF

Def

$f(x)$ is a PDF of random variable X if:

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

CDF

$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$

Expected value

$E[X] = \int_{-\infty}^{\infty} xf(x)dx$

From CDF to PDF

$f(x) = \frac{d}{dx}F(x)$

$f(x)$ must be continuous

3 Convolution

Discrete

X and Y are independent random variables, then $Z = X + Y$ is a random variable with PMF:

$$P(Z = z) = \sum_{x=-\infty}^{\infty} P(X = x)P(Y = z - x)$$

Continuous

X and Y are independent random variables, then $Z = X + Y$ is a random variable with PDF:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z - x)dx$$

4 Gaussian Integral

1. $\int_{-\infty}^{\infty} e^{-ax^2}dx = \sqrt{\frac{\pi}{a}}$
2. $\int_{-\infty}^{\infty} xe^{-a(x-b)^2}dx = b\sqrt{\frac{\pi}{a}}$
3. $\int_{-\infty}^{\infty} x^2e^{-ax^2}dx = \frac{\sqrt{\pi}}{2a^{3/2}}$

5 Joint PDFs and CDFs

Joint Distributions

The **joint CDF** of X and Y is

$$F(x, y) = P(X \leq x, Y \leq y)$$

In the discrete case, X and Y have a **joint PMF**

$$p_{X,Y}(x, y) = P(X = x, Y = y).$$

In the continuous case, they have a **joint PDF**

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y}F_{X,Y}(x, y).$$

The joint PMF/PDF must be nonnegative and sum/integrate to 1.

Marginal Distributions

To find the distribution of one (or more) random variables from a joint PMF/PDF, sum/integrate over the unwanted random variables.

Marginal PMF from joint PMF

$$P(X = x) = \sum_y P(X = x, Y = y)$$

Marginal PDF from joint PDF

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$$

Independence of Random Variables

Random variables X and Y are independent if and only if any of the following conditions holds:

- Joint CDF is the product of the marginal CDFs
- Joint PMF/PDF is the product of the marginal PMFs/PDFs
- Conditional distribution of Y given X is the marginal distribution of Y

6 Linear Transformation of 2 RVs

Let U_1, U_2, V_1, V_2 be r.v.s that satisfy the following conditions:

$V_1 = aU_1 + bU_2$

$V_2 = cU_1 + dU_2$.

Define the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Then we have:

$f_{V_1,V_2}(v_1, v_2) = \frac{1}{|A|} f_{U_1,U_2}(A^{-1}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix})$

7 Moment Generating Functions

Def

MGF For any random variable X, the function

$$M_X(t) = E(e^{tX})$$

$$\mu_k = E(X^k) = M_X^{(k)}(0)$$

MGF of linear functions

MGF of linear functions If we have $Y = aX + b$

$$M_Y(t) = e^{bt}M_X(at)$$

MGF of sum of independent RVs

If X and Y are independent

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

8 Conditional expectation

Def

Conditional PMF The conditional PMF of X given $Y = y$ is

$$f_{X|Y}(x|y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

Conditional PDF The conditional PDF of X given $Y = y$ is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Conditional Expectation discrete The conditional expectation of X given $Y = y$ is

$$E(X|Y = y) = \sum_x x f_{X|Y}(x|y)$$

Conditional Expectation continuous The conditional expectation of X given $Y = y$ is

$$E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx$$

Conditional Expectation Properties

Independence If X and Y are independent random variables, then

$$E(X|Y = y) = E(X)$$

Law of Iterated Expectation If X and Y are random variables, then

$$E(X) = E(E(X|Y))$$

9 CLT

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with mean μ and variance σ^2.

$$\lim n \to \infty P(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq z) = \Phi(z)$$

converges in distribution to a standard normal random variable.

10 Expected value of two random variables

linearity

$E[a \times g_1(X, Y) + b \times g_2(X, Y)] = E[a \times g_1(X, Y)] + E[b \times g_2(X, Y)]$

Independence

If X and Y are independent random variables, then

$E[XY] = E[X]E[Y]$

The converse is NOT true.

Cauchy-Schwarz inequality

$E[XY]^2 \leq E[X^2]E[Y^2]$

Equality holds if and only if $X = aY + b$ for some constants a and b.

11 Bivariate Normal Distribution

Def

X, Y are bivariate normal if X, Y are jointly normal.

X, Y are jointly normal if $\forall a, b \in \mathbb{R}$, $aX + bY$ is normal.

Properties

1. X, Y are jointly normal $\iff$ X, Y are independent.
2. X, Y are jointly normal $\iff$ X, Y are uncorrelated.

joint PDF

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right)$$

joint PDF matrix form

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

where $\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$

construction of bivariate normal r.v.

Let Z_1, Z_2 be independent standard normal r.v.

Let $X = \mu_x + \sigma_x Z_1$, $Y = \mu_y + \sigma_y(\rho Z_1 + \sqrt{1-\rho^2}Z_2)$

Then X, Y are bivariate normal r.v. with $\mu_x, \mu_y, \sigma_x, \sigma_y, \rho$.

12 Covariance

Def

$Cov(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$

Correlation Coefficient

$\rho(X, Y) = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}}$

Properties

- $(Cov(X, Y))^2 \leq Var[X]Var[Y]$
- $Cov(aX, aY) = a^2Cov(X, Y)$
- $-1 \leq \rho(X, Y) \leq 1$
- $\rho(X, Y) = 0 \leftrightarrow independence$
- X_1, X_2 normal does NOT imply X_1, X_2 bivariate normal
  $X_1 = |Y|sign(Z)$
  $X_2 = Y$
  Y, Z are independent standard normal r.v.

13 Concentrarion Inequality

Markov's Inequality

Let X be a non-negative random variable, then for any $t > 0$,

$P(X \geq t) \leq \frac{E[X]}{t}$

Chebyshev's Inequality

Let X be a random variable with finite mean μ and variance σ^2, then for any $t > 0$,

$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$

Chernoff Bound

Let X be a random variable, then for any $t > 0, t \in S, a \in R$

$P(X \geq a) \leq e^{-ta} \times M_x(t)$

Optimize Chernoff Bound

$P(X \geq a) \leq e^{-\phi(a)}$

$\phi(a) = max(ta - ln(M_x(t))), t > 0, t \in S$

Hoeffding's Lemma

Let Z be a r.v. with $E[Z] = 0$, $Z \in [a, b]$ with probability 1

$E[e^{tZ}] \leq exp(\frac{t^2(b-a)^2}{8})$

Hoeffding's Inequality

Let $X_1, ..., X_n$ be a sequence of i.i.d r.v. with $a \leq X_i \leq b$

Define $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$P(|\overline{X} - \mu| \geq \epsilon) \leq 2exp(-\frac{2n\epsilon^2}{(b-a)^2})$

$P(\overline{X} - \mu \geq \epsilon) \leq exp(-\frac{2n\epsilon^2}{(b-a)^2})$

$P(\overline{X} - \mu \leq -\epsilon) \leq exp(-\frac{2n\epsilon^2}{(b-a)^2})$

14 WLLN

Def

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 < \infty$. Then, for any $\epsilon > 0$,

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| > \epsilon) = 0$$

where $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Proof

$P(|\overline{X}_n - \mu| > \epsilon) = P(|\overline{X}_n - \mu|^2 > \epsilon^2)$

By Chebyshev's inequality, we have

$$P(|\overline{X}_n - \mu|^2 > \epsilon^2) \leq \frac{\sigma^2}{n\epsilon^2}$$

Since $\frac{\sigma^2}{n\epsilon^2} \to 0$ as $n \to \infty$, we can conclude that

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu|^2 > \epsilon^2) = 0$$

Therefore, $\overline{X}_n \xrightarrow{p} \mu$.

15 SLLN

Def

The Strong Law of Large Numbers states that if $X_1, X_2, \ldots, X_n$ are i.i.d r.v. with finite mean μ, then the sample mean $\overline{X}_n$ converges almost surely to μ as n approaches infinity, i.e.,

$$P\left(\lim_{n\to\infty} \frac{X_1 + X_2 + \ldots + X_n}{n} = \mu\right) = 1$$

WLLN does not imply SLLN. SLLN implies WLLN.

table of distributions

Distribution	PMF/PDF and Support	Expected Value	Variance	MGF
Bernoulli Bern(p)	$P(X=1) = p$ $P(X=0) = q = 1-p$	p	pq	$q + pe^t$
Binomial Bin(n, p)	$P(X=k) = \binom{n}{k}p^k q^{n-k}$ $k \in \{0, 1, 2, \ldots n\}$	np	npq	$(q+pe^t)^n$
Geometric Geom(p)	$P(X=k) = (1-p)^{k-1}p$ $k \in \{1, 2, \ldots\}$	$1/p$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$, $t < -ln(1-p)$
Poisson Pois(λ, T)	$P(X=k) = \frac{e^{-\lambda T}(\lambda T)^k}{k!}$ $k \in \{0, 1, 2, \ldots\}$	λT	λT	$e^{\lambda T(e^t - 1)}$
Uniform Unif(a, b)	$f(x) = \frac{1}{b-a}$ $x \in (a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$
Discrete Uniform Unif(a, b)	$f(x) = \frac{1}{b-a+1}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2 - 1}{12}$	$\frac{e^{at} - e^{(b+1)t}}{(b-a+1)(1-e^t)}$
Normal $\mathcal{N}(\mu, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}$ $x \in (-\infty, \infty)$	μ	σ^2	$e^{t\mu + \frac{\sigma^2 t^2}{2}}$
Exponential Expo(λ)	$f(x) = \lambda e^{-\lambda x}$ $x \in (0, \infty)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t}$, $t < \lambda$