Sums of independent r.v., covariance and correlation

Proposition (Discrete case) Let X, Y be discrete independent random variables and $Z = X + Y$, then the PMF of Z is

$$p_Z(z) = \sum_x p_X(x) p_Y(z - x).$$

Proposition (Continuous case) Let X, Y be continuous independent random variables and $Z = X + Y$, then the PDF of Z is

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

Proposition (Sum of independent normal r.v.) Let $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ independent. Then $Z = X + Y \sim \mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$.

Definition (Covariance) We define the covariance of random variables X, Y as

$$\operatorname{Cov}(X, Y) \triangleq \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right].$$

Properties (Properties of covariance)

- If X, Y are independent, then $\operatorname{Cov}(X, Y) = 0$.
- $\operatorname{Cov}(X, X) = \operatorname{Var}(X)$.
- $\operatorname{Cov}(aX + b, Y) = a \operatorname{Cov}(X, Y)$.
- $\operatorname{Cov}(X, Y + Z) = \operatorname{Cov}(X, Y) + \operatorname{Cov}(X, Z)$.
- $\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

Proposition (Variance of a sum of r.v.)

$$\operatorname{Var}(X_1 + \cdots + X_n) = \sum_i \operatorname{Var}(X_i) + \sum_{i \neq j} \operatorname{Cov}(X_i, X_j).$$

Definition (Correlation coefficient) We define the correlation coefficient of random variables X, Y, with $\sigma_X, \sigma_Y > 0$, as

$$\rho(X, Y) \triangleq \frac{\operatorname{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

Properties (Properties of the correlation coefficient)

- $-1 \leq \rho \leq 1$.
- If X, Y are independent, then $\rho = 0$.
- $|\rho| = 1$ if and only if $X - \mathbb{E}[X] = c(Y - \mathbb{E}[Y])$.
- $\rho(aX + b, Y) = \operatorname{sign}(a) \rho(X, Y)$.

Conditional expectation and variance, sum of random number of r.v.

Definition (Conditional expectation as a random variable) Given random variables X, Y the conditional expectation $\mathbb{E}[X|Y]$ is the random variable that takes the value $\mathbb{E}[X|Y = y]$ whenever $Y = y$.

Theorem (Law of iterated expectations)

$$\mathbb{E}\left[\mathbb{E}[X|Y]\right] = \mathbb{E}[X].$$

Definition (Conditional variance as a random variable) Given random variables X, Y the conditional variance $\operatorname{Var}(X|Y)$ is the random variable that takes the value $\operatorname{Var}(X|Y = y)$ whenever $Y = y$.

Theorem (Law of total variance)

$$\operatorname{Var}(X) = \mathbb{E}\left[\operatorname{Var}(X|Y)\right] + \operatorname{Var}\left(\mathbb{E}[X|Y]\right).$$

Proposition (Sum of a random number of independent r.v.)
Let N be a nonnegative integer random variable.
Let $X, X_1, X_2, \ldots, X_N$ be i.i.d. random variables.
Let $Y = \sum_i X_i$. Then

$$\mathbb{E}[Y] = \mathbb{E}[N]\mathbb{E}[X],$$
$$\operatorname{Var}(Y) = \mathbb{E}[N] \operatorname{Var}(X) + (\mathbb{E}[X])^2 \operatorname{Var}(N).$$

Convergence of random variables

Inequalities, convergence, and the Weak Law of Large Numbers

Theorem (Markov inequality) Given a random variable $X \geq 0$ and, for every $a > 0$ we have

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

Theorem (Chebyshev inequality) Given a random variable X with $\mathbb{E}[X] = \mu$ and $\operatorname{Var}(X) = \sigma^2$, for every $\epsilon > 0$ we have

$$\mathbb{P}\left(|X - \mu| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}.$$

Theorem (Weak Law of Large Number (WLLN)) Given a sequence of i.i.d. random variables $\{X_1, X_2, \ldots\}$ with $\mathbb{E}[X_i] = \mu$ and $\operatorname{Var}(X_i) = \sigma^2$, we define

$$M_n = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

for every $\epsilon > 0$ we have

$$\lim_{n \to \infty} \mathbb{P}\left(|M_n - \mu| \geq \epsilon\right) = 0.$$

Definition (Convergence in probability) A sequence of random variables $\{Y_i\}$ converges in probability to the random variable Y if

$$\lim_{n \to \infty} \mathbb{P}\left(|Y_i - Y| \geq \epsilon\right) = 0,$$

for every $\epsilon > 0$.

Properties (Properties of convergence in probability) If $X_n \to a$ and $Y_n \to b$ in probability, then

- $X_n + Y_n \to a + b$.
- If g is a continuous function, then $g(X_n) \to g(a)$.
- $\mathbb{E}[X_n]$ does not always converge to a.

The Central Limit Theorem

Theorem (Central Limit Theorem (CLT)) Given a sequence of independent random variables $\{X_1, X_2, \ldots\}$ with $\mathbb{E}[X_i] = \mu$ and $\operatorname{Var}(X_i) = \sigma^2$, we define

$$Z_n = \frac{1}{\sigma \sqrt{n}} \sum_{i=1}^{n} (X_i - \mu).$$

Then, for every z, we have

$$\lim_{n \to \infty} \mathbb{P}(Z_n \leq z) = \mathbb{P}(Z \leq z),$$

where $Z \sim \mathcal{N}(0, 1)$.

Corollary (Normal approximation of a binomial) Let $X \sim Bin(n, p)$ with n large. Then S_n can be approximated by $Z \sim \mathcal{N}\left(np, np(1 - p)\right)$.

Remark (De Moivre-Laplace 1/2 approximation) Let $X \sim Bin$, then $\mathbb{P}(X = i) = \mathbb{P}\left(i - \frac{1}{2} \leq X \leq i + \frac{1}{2}\right)$ and we can use the CLT to approximate the PMF of X.