MATH 3338 Probability

Wenjiang Fu wfu2@central.uh.edu

Lecture 7 - MATH 3338
Ch 7 Sums of Independent Random Variables

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Outline

1. Sums of Discrete Random Variables
2. Sums of Continuous Random Variables

Sums of Discrete Random Variables

We consider the sum of independent random variables and distribution.

Convolution

Suppose X and Y are two indep discrete RVs with distributions $m_1(x)$ and $m_2(x)$. Let $Z = X + Y$. We need to determine the distribution function $m_3(x)$ of Z. Suppose $X = k$, some integer. Then $Z = z$ iff $Y = z - k$. So the event $Z = z$ is the union of the pairwise disjoint events

$$(X = k) \cap (Y = z - k),$$

where k runs over the integers. Since the events of different k are pairwise disjoint

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

This leads to the distribution function of Z, and the definition below.

Sums of Discrete Random Variables

Convolution Definition 7.1 Let X and Y be two indep integer-valued RVs with distributions $m_1(x)$ and $m_2(x)$, respectively. Then the convolution of $m_1(x)$ and $m_2(x)$ is the distribution function $m_3 = m_1 * m_2$ given by

$$m_3(j) = \sum_k m_1(k) \cdot m_2(j - k),$$

for $j = ..., -2, -1, 0, 1, 2, ...$ The function $m_3(x)$ is the distribution function of the RV $Z = X + Y$.

Special cases of convolution where we did not even think about convolution.

1) Sum of indep. Bernoulli trials results in Binomial.

2) Sum of indep. geometric results in negative binomial.

3) Sum of Chi-squares results in Chi-squares with more degrees of freedom.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Sums of Discrete Random Variables

Convolution It is easy to see that convolution is commutative and associative.
Let $S_n = X_1 + X_2 + ... + X_n$ be the sum of n indep trials process with common distribution function m defined on the integers. Then the distribution function of S_1 is m.

$$S_n = S_{n-1} + X_n.$$

The distribution function of S_n can be found by induction.
Example 7.1 A die is rolled twice. Let X_1, X_2 be the outcomes, and let $S_2 = X_1 + X_2$ be the sum of these outcomes. Then X_1 and X_2 have the common distribution function: $m(i) = 1/6$ for all $i = 1, ..., 6$.
The distribution function of S_2 is then the convolution of this distribution with itself.

Sums of Discrete Random Variables

Convolution $P(S_2 = 2) = m(1)m(1) = 1/6^2 = 1/36$
$P(S_2 = 3) = m(1)m(2) + m(2)m(1) = 1/6^2 + 1/6^2 = 1/36$
$P(S_2 = 4) = m(1)m(3) + m(3)m(1) + m(2)m(2) = 3/36$.
Continuing in this way, we find $P(S_2 = 5) = 4/36$, $P(S_2 = 6) = 5/36$, $P(S_2 = 7) = 6/36$, $P(S_2 = 8) = 5/36$, $P(S_2 = 9) = 4/36$, $P(S_2 = 10) = 3/36$, $P(S_2 = 11) = 2/36$, $P(S_2 = 12) = 1/36$.
Furthermore, the distribution for S_3 would be the convolution of the distributio for S_2 with the distribution for S_3.

$$P(S_3) = P(S_2 = 2)P(X_3 = 1) = 1/36 \cdot 1/6 = 1/216.$$

Such tedious job can be programmed and completed by computer.

We consider the sum of independent random variables and distribution.

Convolution

Definition 7.2 Let X and Y be two continuous rs with density functions $f(x)$ and $g(y)$. Assume both of them are defined for all real numbers. Then the convolution $f * g$ of f and g is the function given by

$$(f * g)(z) = \int_{-\infty}^{+\infty} f(z-y)g(y)dy = \int_{-\infty}^{+\infty} g(z-x)f(x)dx$$

Theorem 7.1 Let X and Y be two continuous rs with density functions $f_X(x)$ and $f_Y(y)$ defined for all x. Then the sum $Z = X + Y$ is a rv with density function $f_Z(z)$, where $f_Z(z)$ is the convolution of f_X and g_Y.

Sums of Continuous Random Variables

Examples

- Sum of Two Uniform Random Variables.
  Let $X, Y \sim Unif[0,1]$ are indep. $Z = X + Y$.

$$f_Z(z) = \int_0^1 f_X(z-y)f_Y(y)dy = \int_{0\le z-y\le 1, 0\le y\le 1} 1dy$$

$$= \int_{z-1\le y\le z, 0\le y\le 1} 1dy = \begin{cases} z, & 0\le z\le 1 \\ 2-z, & 1\le z\le 2 \\ 0, & \text{otherwise} \end{cases}$$

- Sum of Two Exponential Random Variables.

$$X, Y \sim exp(\lambda) : f(x) = \lambda e^{-\lambda x}, \forall x \ge 0.$$

$$f(z) = \int_0^{+\infty} \lambda e^{-\lambda(z-y)} \lambda e^{-\lambda y} dy = \lambda^2 \int_0^z e^{-\lambda z} dz = \lambda^2 z e^{-\lambda z}, \quad \forall z \ge 0$$

Sums of Continuous Random Variables

Examples of Sums of Continuous RVs.

- **Sums of Indep Normal RVs.**
  Let $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$. Since
  $E(X+Y) = E(X) + E(Y) = \mu_1 + \mu_2$, and
  $Var(X+Y) = Var(X) + Var(Y) = \sigma_1^2 + \sigma_2^2$. We simplify the convolution to consider the sum of two RVs $X, Y \sim N(0,1)$.
  $Z = X + Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{(\sqrt{2\pi})^2} e^{-(z-y)^2/2} e^{-y^2/2} dy = \frac{1}{2\pi} e^{-z^2/2} \int_{-\infty}^{\infty} e^{-y^2-yz} dy$$

$$= \frac{1}{\sqrt{4\pi}} e^{-y^2/4}$$

- $$Z = X + Y \sim N(0,2)$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Examples of Sums of Continuous RVs.

- **Sums of Indep Cauchy RVs.**
  Let $X, Y \sim Cauchy(a), a = 1$, indep. $Z = X + Y$.

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2}$$

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\pi^2} \frac{1}{1 + (z - y)^2} \frac{1}{1 + y^2} dy = \frac{2}{\pi(4 + z^2)}$$

If $Z = (X + Y)/2$, then

$$f_Z(z) = \frac{1}{\pi(1 + z^2)}$$

going back to Cauchy with parameter $a = 1$.