Math 461 Fall 2021

Renming Song

University of Illinois at Urbana-Champaign

October 22, 2021

Outline

Outline

1 General Info

2 6.3 Sums of independent random variables

HW7 is due today before the end of class time . Please submit your HW7 via the course Moodle page.

Solution to HW7 will be on my homepage this weekend.

HW7 is due today before the end of class time . Please submit your HW7 via the course Moodle page.

Solution to HW7 will be on my homepage this weekend.

Outline

1 General Info

2 6.3 Sums of independent random variables

Last time, we have seen that, if X and Y are independent abs. cont. random variables with density f_X and f_Y respectively, then the density of $Z = X + Y$ is

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

We also have

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy.$$

Now let's suppose that X and Y are independent positive abs. cont. random variables with density f_X and f_Y respectively, then $Z = X + Y$ is a also a positive random variable and its density is

Last time, we have seen that, if X and Y are independent abs. cont. random variables with density f_X and f_Y respectively, then the density of $Z = X + Y$ is

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

We also have

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy.$$

Now let's suppose that X and Y are independent positive abs. cont. random variables with density f_X and f_Y respectively, then $Z = X + Y$ is a also a positive random variable and its density is

$$f_Z(z) = \begin{cases} \int_0^z f_X(x) f_Y(z-x)dx, & z > 0, \\ 0, & \text{otherwise.} \end{cases}$$

We also have

$$f_Z(z) = \begin{cases} \int_0^z f_X(z-y) f_Y(y)dy, & z > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition

Suppose X and Y are independent random variables.

(i) If X and Y are Gamma random variables with parameters (α, λ) and (β, λ) respectively, then $X + Y$ is a Gamma random variable with parameters $(\alpha + \beta, \lambda)$.

(ii) If X and Y are normal random variables with parameters (μ_1, σ_1^2) and (μ_2, σ_2^2) respectively, then $X + Y$ is a normal random variable $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

$$f_Z(z) = \begin{cases} \int_0^z f_X(x) f_Y(z-x)dx, & z > 0, \\ 0, & \text{otherwise.} \end{cases}$$

We also have

$$f_Z(z) = \begin{cases} \int_0^z f_X(z-y) f_Y(y)dy, & z > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition

Suppose X and Y are independent random variables.

(i) If X and Y are Gamma random variables with parameters (α, λ) and (β, λ) respectively, then $X + Y$ is a Gamma random variable with parameters $(\alpha + \beta, \lambda)$.

(ii) If X and Y are normal random variables with parameters (μ_1, σ_1^2) and (μ_2, σ_2^2) respectively, then $X + Y$ is a normal random variable $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Let's prove (i). For any $z > 0$,

$$
\begin{aligned}
f_{X+Y}(z) &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^z \lambda e^{-\lambda x}(\lambda x)^{\alpha-1} \lambda e^{-\lambda(z-x)}(\lambda(z-x))^{\beta-1} dx \\
&= \frac{\lambda e^{-\lambda z}}{\Gamma(\alpha)\Gamma(\beta)} \lambda^{\alpha+\beta-1} \int_0^z x^{\alpha-1}(z-x)^{\beta-1} dx \\
&= \frac{\lambda e^{-\lambda z}}{\Gamma(\alpha)\Gamma(\beta)} (\lambda z)^{\alpha+\beta-1} \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du, \quad x = zu, \\
&= \frac{\lambda e^{-\lambda z}}{\Gamma(\alpha)\Gamma(\beta)} (\lambda z)^{\alpha+\beta-1} B(\alpha, \beta) \\
&= \frac{1}{\Gamma(\alpha+\beta)} \lambda e^{-\lambda z} (\lambda z)^{\alpha+\beta-1}.
\end{aligned}
$$

Example 1

A basketball team will play a 44-game season. 26 of these games are against class A teams and 18 are are against class B teams. Suppose that the team will win each game against a class A team with probability .4 and will win each game against a class B team with probability .7. Suppose also that the results of different games are independent. Approximate the probability that

(a) the team wins 25 or more games;

(b) the team will win more games against class A teams than it does agains class B teams.

Let X_A and X_B denote respectively the number of games the teams wins are against class A teams and are against class B teams. Then X_A and X_B are independent binomial random variables with parameters (26, .4) and (18, .7) respectively.

Example 1

A basketball team will play a 44-game season. 26 of these games are against class A teams and 18 are are against class B teams. Suppose that the team will win each game against a class A team with probability .4 and will win each game against a class B team with probability .7. Suppose also that the results of different games are independent. Approximate the probability that

(a) the team wins 25 or more games;

(b) the team will win more games against class A teams than it does agains class B teams.

Let X_A and X_B denote respectively the number of games the teams wins are against class A teams and are against class B teams. Then X_A and X_B are independent binomial random variables with parameters $(26, .4)$ and $(18, .7)$ respectively.

$$E[X_A] = 26(.4) = 10.4, \quad \mathrm{Var}(X_A) = 26(.4)(.6) = 6.24$$
$$E[X_B] = 18(.7) = 12.6, \quad \mathrm{Var}(X_B) = 18(.7)(.3) = 3.78.$$

By the central limit theorem, X_A is approximately normal with parameters $(10.4, 6.24)$ and X_B is approximately normal with parameters $(12.6, 3.78)$.

By the Proposition above, $X_A + X_B$ is approximately normal with parameters $(23, 10.02)$ since X_A and X_B are independent. Thus

$$\begin{aligned} P(X_A + X_B \geq 25) &= P(X_A + X_B \geq 24.5) \\ &= P\left(\frac{X_A + X_B - 23}{\sqrt{10.02}} \geq \frac{24.5 - 23}{\sqrt{10.02}}\right) \\ &= P\left(\frac{X_A + X_B - 23}{\sqrt{10.02}} \geq .4739\right) \approx 1 - \Phi(.4739) \approx .3178. \end{aligned}$$

$$E[X_A] = 26(.4) = 10.4, \quad \mathrm{Var}(X_A) = 26(.4)(.6) = 6.24$$
$$E[X_B] = 18(.7) = 12.6, \quad \mathrm{Var}(X_B) = 18(.7)(.3) = 3.78.$$

By the central limit theorem, X_A is approximately normal with parameters $(10.4, 6.24)$ and X_B is approximately normal with parameters $(12.6, 3.78)$.

By the Proposition above, $X_A + X_B$ is approximately normal with parameters $(23, 10.02)$ since X_A and X_B are independent. Thus

$$P(X_A + X_B \geq 25) = P(X_A + X_B \geq 24.5)$$
$$= P\left(\frac{X_A + X_B - 23}{\sqrt{10.02}} \geq \frac{24.5 - 23}{\sqrt{10.02}}\right)$$
$$= P\left(\frac{X_A + X_B - 23}{\sqrt{10.02}} \geq .4739\right) \approx 1 - \Phi(.4739) \approx .3178.$$

Since X_A and X_B are independent, by the Proposition above, $X_A - X_B$ is approximately normal with parameters $(-2.2, 10.02)$. Hence

$$
\begin{aligned}
&P(X_A - X_B \geq 1) = P(X_A - X_B \geq .5) \\
&= P\left(\frac{X_A - X_B + 2.2}{\sqrt{10.02}} \geq \frac{.5 + 2.2}{\sqrt{10.02}}\right) \\
&= P\left(\frac{X_A - X_B + 2.2}{\sqrt{10.02}} \geq .8530\right) \approx 1 - \Phi(.8530) \approx .1968.
\end{aligned}
$$

Example 2

Suppose that X and Y are independent standard normal random variables. Find the density of $Z = X^2 + Y^2$.

Since X_A and X_B are independent, by the Proposition above, $X_A - X_B$ is approximately normal with parameters $(-2.2, 10.02)$. Hence

$$P(X_A - X_B \geq 1) = P(X_A - X_B \geq .5)$$
$$= P\left(\frac{X_A - X_B + 2.2}{\sqrt{10.02}} \geq \frac{.5 + 2.2}{\sqrt{10.02}}\right)$$
$$= P\left(\frac{X_A - X_B + 2.2}{\sqrt{10.02}} \geq .8530\right) \approx 1 - \Phi(.8530) \approx .1968.$$

Example 2

Suppose that X and Y are independent standard normal random variables. Find the density of $Z = X^2 + Y^2$.

We know that X^2 and Y^2 are independent Gamma random variables with parameters $(\frac{1}{2}, \frac{1}{2})$. Thus $X^2 + Y^2$ is a Gamma random variables with parameters $(1, \frac{1}{2})$, that is, an exponential random variable with parameter $1/2$.

Example 3

Suppose that X and Y are independent random variables, both uniformly distributed on $(0, 1)$. Find the density of $Z = X + Y$.

Applying the formula directly is not easy. We look for the distribution of Z first.

We know that X^2 and Y^2 are independent Gamma random variables with parameters $(\frac{1}{2}, \frac{1}{2})$. Thus $X^2 + Y^2$ is a Gamma random variables with parameters $(1, \frac{1}{2})$, that is, an exponential random variable with parameter $1/2$.

Example 3

Suppose that X and Y are independent random variables, both uniformly distributed on $(0, 1)$. Find the density of $Z = X + Y$.

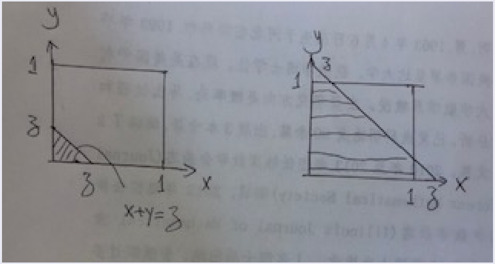
Applying the formula directly is not easy. We look for the distribution of Z first.

We know that X^2 and Y^2 are independent Gamma random variables with parameters $(\frac{1}{2}, \frac{1}{2})$. Thus $X^2 + Y^2$ is a Gamma random variables with parameters $(1, \frac{1}{2})$, that is, an exponential random variable with parameter $1/2$.

Example 3

Suppose that X and Y are independent random variables, both uniformly distributed on $(0, 1)$. Find the density of $Z = X + Y$.

Applying the formula directly is not easy. We look for the distribution of Z first.

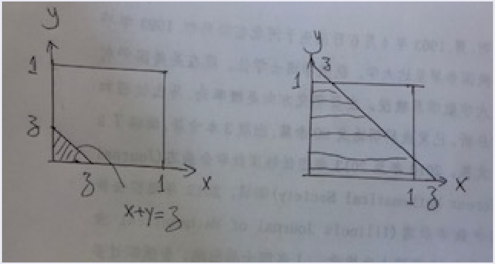


$X + Y$ takes values in $(0, 2)$. For $z \in (0, 1]$,

$$P(Z \le z) = P(X + Y \le z) = \frac{z^2}{2}.$$

For $z \in (1, 2)$,

$$P(Z \le z) = P(X + Y \le z) = 1 - \frac{(2 - z)^2}{2}.$$



$X + Y$ takes values in $(0, 2)$. For $z \in (0, 1]$,

$$P(Z \leq z) = P(X + Y \leq z) = \frac{z^2}{2}.$$

For $z \in (1, 2)$,

$$P(Z \leq z) = P(X + Y \leq z) = 1 - \frac{(2 - z)^2}{2}.$$

Thus the density of Z is

$$f_Z(z) = \begin{cases} z, & 0 \leq z \leq 1, \\ 2 - z, & 1 < z < 2, \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that X and Y are independent discrete random variables with mass functions $p_X(\cdot)$ and $p_Y(\cdot)$ respectively. Find the mass function of $Z = X + Y$.

Thus the density of Z is

$$f_Z(z) = \begin{cases} z, & 0 \leq z \leq 1, \\ 2 - z, & 1 < z < 2, \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that X and Y are independent discrete random variables with mass functions $p_X(\cdot)$ and $p_Y(\cdot)$ respectively. Find the mass function of $Z = X + Y$.

For any z,

$$p_Z(z) = P(X+Y=z) = \sum_x P(X+Y=z, X=x)$$
$$= \sum_x P(X=x, Y=z-x) = \sum_x P(X=x)P(Y=z-x)$$
$$= \sum_x p_X(x)p_Y(z-x).$$

We also have

$$p_Z(z) = \sum_y p_X(z-y)p_Y(y).$$

For any z,

$$p_Z(z) = P(X+Y=z) = \sum_x P(X+Y=z, X=x)$$
$$= \sum_x P(X=x, Y=z-x) = \sum_x P(X=x)P(Y=z-x)$$
$$= \sum_x p_X(x)p_Y(z-x).$$

We also have

$$p_Z(z) = \sum_y p_X(z-y)p_Y(y).$$

If X and Y are integer-valued, then for any integer z,

$$p_{X+Y}(z) = \sum_{x=-\infty}^{\infty} p_X(x) p_Y(z-x).$$

If X and Y are non-negative integer-valued, then for any non-negative integer z,

$$p_{X+Y}(z) = \sum_{x=0}^{z} p_X(x) p_Y(z-x).$$

If X and Y are positive integer-valued, then $X+Y$ takes values $2, 3, \ldots$. For $z = 2, 3, \ldots,$

$$p_{X+Y}(z) = \sum_{x=1}^{z-1} p_X(x) p_Y(z-x).$$

If X and Y are integer-valued, then for any integer z,

$$p_{X+Y}(z) = \sum_{x=-\infty}^{\infty} p_X(x) p_Y(z-x).$$

If X and Y are non-negative integer-valued, then for any non-negative integer z,

$$p_{X+Y}(z) = \sum_{x=0}^{z} p_X(x) p_Y(z-x).$$

If X and Y are positive integer-valued, then $X+Y$ takes values $2, 3, \ldots$. For $z = 2, 3, \ldots$,

$$p_{X+Y}(z) = \sum_{x=1}^{z-1} p_X(x) p_Y(z-x).$$

Proposition

Suppose that X and Y are independent random variables.

(i) If X is a binomial random variable with parameters (m, p), and Y is a binomial random variable with parameters (n, p), then $X + Y$ is a binomial random variable with parameters $(m + n, p)$;

(ii) If X is a Poisson random variables with parameter λ_1, and Y is a Poisson random variables with parameter λ_2, then $X + Y$ is a Poisson random variables with parameter $\lambda_1 + \lambda_2$;

(iii) If X is a negative binomial random variable with parameters (r_1, p), and Y is a negative binomial random variable with parameters (r_2, p), then $X + Y$ is a negative binomial random variable with parameters $(r_1 + r_2, p)$.

I will only give the proof of (ii).

Proposition

Suppose that X and Y are independent random variables.

(i) If X is a binomial random variable with parameters (m, p), and Y is a binomial random variable with parameters (n, p), then $X + Y$ is a binomial random variable with parameters $(m + n, p)$;

(ii) If X is a Poisson random variables with parameter λ_1, and Y is a Poisson random variables with parameter λ_2, then $X + Y$ is a Poisson random variables with parameter $\lambda_1 + \lambda_2$;

(iii) If X is a negative binomial random variable with parameters (r_1, p), and Y is a negative binomial random variable with parameters (r_2, p), then $X + Y$ is a negative binomial random variable with parameters $(r_1 + r_2, p)$.

I will only give the proof of (ii).

For any $z = 0, 1, \ldots,$

$$p_{X+Y}(z) = \sum_{x=0}^{z} e^{-\lambda_1} \frac{\lambda_1^x}{x!} e^{-\lambda_2} \frac{\lambda_2^{z-x}}{(z-x)!}$$
$$= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^z}{z!} \sum_{x=0}^{z} \binom{z}{x} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{z-x}$$
$$= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^z}{z!}.$$

Example 4

Suppose that X and Y are independent geometric random variables with a common parameter p. Find (a) the mass function of $\min(X, Y)$; (b) $P(\min(X, Y) = X) = P(Y \geq X)$.

For any $z = 0, 1, \ldots,$

$$p_{X+Y}(z) = \sum_{x=0}^{z} e^{-\lambda_1} \frac{\lambda_1^x}{x!} e^{-\lambda_2} \frac{\lambda_2^{z-x}}{(z-x)!}$$
$$= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^z}{z!} \sum_{x=0}^{z} \binom{z}{x} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{z-x}$$
$$= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^z}{z!}.$$

Example 4

Suppose that X and Y are independent geometric random variables with a common parameter p. Find (a) the mass function of $\min(X, Y)$; (b) $P(\min(X, Y) = X) = P(Y \geq X)$.

$\min(X, Y)$ takes only positive integer values. For $z = 1, 2, \ldots,$

$$P(\min(X, Y) > z) = P(X > z, Y > z) = P(X > z)P(Y > z)$$
$$= (1-p)^{2z} = (1 - (2p - p^2))^z.$$

Thus $\min(X, Y)$ is a geometric random variable with parameter $2p - p^2$.

$$P(Y \geq X) = \sum_{x=1}^{\infty} P(X = x, Y \geq X) = \sum_{x=1}^{\infty} P(X = x, Y \geq x)$$
$$= \sum_{x=1}^{\infty} P(X = x)P(Y \geq x) = \sum_{x=1}^{\infty} p(1-p)^{x-1}(1-p)^{x-1}$$
$$= p\sum_{x=1}^{\infty}(1 - (2p - p^2))^{x-1} = \frac{p}{2p - p^2} = \frac{1}{2 - p}.$$

$\min(X, Y)$ takes only positive integer values. For $z = 1, 2, \ldots,$

$$P(\min(X, Y) > z) = P(X > z, Y > z) = P(X > z)P(Y > z)$$
$$= (1-p)^{2z} = (1-(2p-p^2))^z.$$

Thus $\min(X, Y)$ is a geometric random variable with parameter $2p - p^2$.

$$P(Y \geq X) = \sum_{x=1}^{\infty} P(X = x, Y \geq X) = \sum_{x=1}^{\infty} P(X = x, Y \geq x)$$
$$= \sum_{x=1}^{\infty} P(X = x)P(Y \geq x) = \sum_{x=1}^{\infty} p(1-p)^{x-1}(1-p)^{x-1}$$
$$= p\sum_{x=1}^{\infty}(1-(2p-p^2))^{x-1} = \frac{p}{2p-p^2} = \frac{1}{2-p}.$$

Suppose that X and Y are independent random variables such that

$$P(X = i) = P(Y = i) = \frac{1}{100}, \quad i = 1, \dots 100.$$

Find (a) $P(X \geq Y)$; (b) $P(X = Y)$.

$$P(X \geq Y) = \sum_{y=1}^{100} P(X \geq Y, Y = y) = \sum_{y=1}^{100} P(X \geq y)P(Y = y)$$

$$= \frac{1}{100^2} \sum_{y=1}^{100} (101 - y) = \frac{1}{100^2} \sum_{i=1}^{100} i = \frac{101}{200}.$$

$$P(X = Y) = \sum_{y=1}^{100} P(X = x, Y = X) = \sum_{y=1}^{100} P(X = x, Y = x)$$

$$= \sum_{y=1}^{100} P(X = x)P(Y = x) = \frac{1}{100}.$$

Suppose that X and Y are independent random variables such that

$$P(X = i) = P(Y = i) = \frac{1}{100}, \quad i = 1, \ldots 100.$$

Find (a) $P(X \geq Y)$; (b) $P(X = Y)$.

$$P(X \geq Y) = \sum_{y=1}^{100} P(X \geq Y, Y = y) = \sum_{y=1}^{100} P(X \geq y) P(Y = y)$$

$$= \frac{1}{100^2} \sum_{y=1}^{100} (101 - y) = \frac{1}{100^2} \sum_{i=1}^{100} i = \frac{101}{200}.$$

$$P(X = Y) = \sum_{y=1}^{100} P(X = x, Y = X) = \sum_{y=1}^{100} P(X = x, Y = x)$$

$$= \sum_{y=1}^{100} P(X = x) P(Y = x) = \frac{1}{100}.$$