

18.600: Lecture 22

Sums of independent random variables

Scott Sheffield

MIT

Summing two random variables

- ▶ Say we have independent random variables X and Y and we know their density functions f_X and f_Y .

Summing two random variables

- ▶ Say we have independent random variables X and Y and we know their density functions f_X and f_Y .
- ▶ Now let's try to find $F_{X+Y}(a) = P\{X + Y \leq a\}$.

Summing two random variables

- ▶ Say we have independent random variables X and Y and we know their density functions f_X and f_Y .
- ▶ Now let's try to find $F_{X+Y}(a) = P\{X + Y \leq a\}$.
- ▶ This is the integral over $\{(x, y) : x + y \leq a\}$ of $f(x, y) = f_X(x)f_Y(y)$. Thus,

Summing two random variables

- ▶ Say we have independent random variables X and Y and we know their density functions f_X and f_Y .
- ▶ Now let's try to find $F_{X+Y}(a) = P\{X + Y \leq a\}$.
- ▶ This is the integral over $\{(x, y) : x + y \leq a\}$ of $f(x, y) = f_X(x)f_Y(y)$. Thus,



$$\begin{aligned}P\{X + Y \leq a\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)f_Y(y)dx dy \\ &= \int_{-\infty}^{\infty} F_X(a - y)f_Y(y)dy.\end{aligned}$$

Summing two random variables

- ▶ Say we have independent random variables X and Y and we know their density functions f_X and f_Y .
- ▶ Now let's try to find $F_{X+Y}(a) = P\{X + Y \leq a\}$.
- ▶ This is the integral over $\{(x, y) : x + y \leq a\}$ of $f(x, y) = f_X(x)f_Y(y)$. Thus,



$$\begin{aligned}P\{X + Y \leq a\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)f_Y(y)dx dy \\ &= \int_{-\infty}^{\infty} F_X(a-y)f_Y(y)dy.\end{aligned}$$

- ▶ Differentiating both sides gives $f_{X+Y}(a) = \frac{d}{da} \int_{-\infty}^{\infty} F_X(a-y)f_Y(y)dy = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$.

Summing two random variables

- ▶ Say we have independent random variables X and Y and we know their density functions f_X and f_Y .
- ▶ Now let's try to find $F_{X+Y}(a) = P\{X + Y \leq a\}$.
- ▶ This is the integral over $\{(x, y) : x + y \leq a\}$ of $f(x, y) = f_X(x)f_Y(y)$. Thus,



$$\begin{aligned}P\{X + Y \leq a\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)f_Y(y)dx dy \\ &= \int_{-\infty}^{\infty} F_X(a-y)f_Y(y)dy.\end{aligned}$$

- ▶ Differentiating both sides gives $f_{X+Y}(a) = \frac{d}{da} \int_{-\infty}^{\infty} F_X(a-y)f_Y(y)dy = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$.
- ▶ Latter formula makes some intuitive sense. We're integrating over the set of x, y pairs that add up to a .

Independent identically distributed (i.i.d.)

- ▶ The abbreviation i.i.d. means independent identically distributed.

Independent identically distributed (i.i.d.)

- ▶ The abbreviation i.i.d. means independent identically distributed.
- ▶ It is actually one of the most important abbreviations in probability theory.

Independent identically distributed (i.i.d.)

- ▶ The abbreviation i.i.d. means independent identically distributed.
- ▶ It is actually one of the most important abbreviations in probability theory.
- ▶ Worth memorizing.

Summing i.i.d. uniform random variables

- ▶ Suppose that X and Y are i.i.d. and uniform on $[0, 1]$. So $f_X = f_Y = 1$ on $[0, 1]$.

Summing i.i.d. uniform random variables

- ▶ Suppose that X and Y are i.i.d. and uniform on $[0, 1]$. So $f_X = f_Y = 1$ on $[0, 1]$.
- ▶ What is the probability density function of $X + Y$?

Summing i.i.d. uniform random variables

- ▶ Suppose that X and Y are i.i.d. and uniform on $[0, 1]$. So $f_X = f_Y = 1$ on $[0, 1]$.
- ▶ What is the probability density function of $X + Y$?
- ▶ $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy = \int_0^1 f_X(a-y)$ which is the length of $[0, 1] \cap [a-1, a]$.

Summing i.i.d. uniform random variables

- ▶ Suppose that X and Y are i.i.d. and uniform on $[0, 1]$. So $f_X = f_Y = 1$ on $[0, 1]$.
- ▶ What is the probability density function of $X + Y$?
- ▶ $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy = \int_0^1 f_X(a-y)$ which is the length of $[0, 1] \cap [a-1, a]$.
- ▶ That's a when $a \in [0, 1]$ and $2-a$ when $a \in [1, 2]$ and 0 otherwise.

Review: summing i.i.d. geometric random variables

- ▶ A geometric random variable X with parameter p has $P\{X = k\} = (1 - p)^{k-1}p$ for $k \geq 1$.

Review: summing i.i.d. geometric random variables

- ▶ A geometric random variable X with parameter p has $P\{X = k\} = (1 - p)^{k-1}p$ for $k \geq 1$.
- ▶ Sum Z of n independent copies of X ?

Review: summing i.i.d. geometric random variables

- ▶ A geometric random variable X with parameter p has $P\{X = k\} = (1 - p)^{k-1}p$ for $k \geq 1$.
- ▶ Sum Z of n independent copies of X ?
- ▶ We can interpret Z as time slot where n th head occurs in i.i.d. sequence of p -coin tosses.

Review: summing i.i.d. geometric random variables

- ▶ A geometric random variable X with parameter p has $P\{X = k\} = (1 - p)^{k-1}p$ for $k \geq 1$.
- ▶ Sum Z of n independent copies of X ?
- ▶ We can interpret Z as time slot where n th head occurs in i.i.d. sequence of p -coin tosses.
- ▶ So Z is negative binomial (n, p) . So $P\{Z = k\} = \binom{k-1}{n-1}p^{n-1}(1 - p)^{k-n}p$.

Summing i.i.d. exponential random variables

- ▶ Suppose X_1, \dots, X_n are i.i.d. exponential random variables with parameter λ . So $f_{X_i}(x) = \lambda e^{-\lambda x}$ on $[0, \infty)$ for all $1 \leq i \leq n$.

Summing i.i.d. exponential random variables

- ▶ Suppose X_1, \dots, X_n are i.i.d. exponential random variables with parameter λ . So $f_{X_i}(x) = \lambda e^{-\lambda x}$ on $[0, \infty)$ for all $1 \leq i \leq n$.
- ▶ What is the law of $Z = \sum_{i=1}^n X_i$?

Summing i.i.d. exponential random variables

- ▶ Suppose X_1, \dots, X_n are i.i.d. exponential random variables with parameter λ . So $f_{X_i}(x) = \lambda e^{-\lambda x}$ on $[0, \infty)$ for all $1 \leq i \leq n$.
- ▶ What is the law of $Z = \sum_{i=1}^n X_i$?
- ▶ We claimed in an earlier lecture that this was a gamma distribution with parameters (λ, n) .

Summing i.i.d. exponential random variables

- ▶ Suppose X_1, \dots, X_n are i.i.d. exponential random variables with parameter λ . So $f_{X_i}(x) = \lambda e^{-\lambda x}$ on $[0, \infty)$ for all $1 \leq i \leq n$.
- ▶ What is the law of $Z = \sum_{i=1}^n X_i$?
- ▶ We claimed in an earlier lecture that this was a gamma distribution with parameters (λ, n) .
- ▶ So $f_Z(y) = \frac{\lambda e^{-\lambda y} (\lambda y)^{n-1}}{\Gamma(n)}$.

Summing i.i.d. exponential random variables

- ▶ Suppose X_1, \dots, X_n are i.i.d. exponential random variables with parameter λ . So $f_{X_i}(x) = \lambda e^{-\lambda x}$ on $[0, \infty)$ for all $1 \leq i \leq n$.
- ▶ What is the law of $Z = \sum_{i=1}^n X_i$?
- ▶ We claimed in an earlier lecture that this was a gamma distribution with parameters (λ, n) .
- ▶ So $f_Z(y) = \frac{\lambda e^{-\lambda y} (\lambda y)^{n-1}}{\Gamma(n)}$.
- ▶ We argued this point by taking limits of negative binomial distributions. Can we check it directly?

Summing i.i.d. exponential random variables

- ▶ Suppose X_1, \dots, X_n are i.i.d. exponential random variables with parameter λ . So $f_{X_i}(x) = \lambda e^{-\lambda x}$ on $[0, \infty)$ for all $1 \leq i \leq n$.
- ▶ What is the law of $Z = \sum_{i=1}^n X_i$?
- ▶ We claimed in an earlier lecture that this was a gamma distribution with parameters (λ, n) .
- ▶ So $f_Z(y) = \frac{\lambda e^{-\lambda y} (\lambda y)^{n-1}}{\Gamma(n)}$.
- ▶ We argued this point by taking limits of negative binomial distributions. Can we check it directly?
- ▶ By induction, would suffice to show that a gamma $(\lambda, 1)$ plus an independent gamma (λ, n) is a gamma $(\lambda, n + 1)$.

Summing independent gamma random variables

- ▶ Say X is gamma (λ, s) , Y is gamma (λ, t) , and X and Y are independent.

Summing independent gamma random variables

- ▶ Say X is gamma (λ, s) , Y is gamma (λ, t) , and X and Y are independent.
- ▶ Intuitively, X is amount of time till we see s events, and Y is amount of subsequent time till we see t more events.

Summing independent gamma random variables

- ▶ Say X is gamma (λ, s) , Y is gamma (λ, t) , and X and Y are independent.
- ▶ Intuitively, X is amount of time till we see s events, and Y is amount of subsequent time till we see t more events.
- ▶ So $f_X(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{s-1}}{\Gamma(s)}$ and $f_Y(y) = \frac{\lambda e^{-\lambda y} (\lambda y)^{t-1}}{\Gamma(t)}$.

Summing independent gamma random variables

- ▶ Say X is gamma (λ, s) , Y is gamma (λ, t) , and X and Y are independent.
- ▶ Intuitively, X is amount of time till we see s events, and Y is amount of subsequent time till we see t more events.
- ▶ So $f_X(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{s-1}}{\Gamma(s)}$ and $f_Y(y) = \frac{\lambda e^{-\lambda y} (\lambda y)^{t-1}}{\Gamma(t)}$.
- ▶ Now $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy$.

Summing independent gamma random variables

- ▶ Say X is gamma (λ, s) , Y is gamma (λ, t) , and X and Y are independent.
- ▶ Intuitively, X is amount of time till we see s events, and Y is amount of subsequent time till we see t more events.
- ▶ So $f_X(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{s-1}}{\Gamma(s)}$ and $f_Y(y) = \frac{\lambda e^{-\lambda y} (\lambda y)^{t-1}}{\Gamma(t)}$.
- ▶ Now $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy$.
- ▶ Up to an a -independent multiplicative constant, this is

$$\int_0^a e^{-\lambda(a-y)} (a-y)^{s-1} e^{-\lambda y} y^{t-1} dy = e^{-\lambda a} \int_0^a (a-y)^{s-1} y^{t-1} dy.$$

Summing independent gamma random variables

- ▶ Say X is gamma (λ, s) , Y is gamma (λ, t) , and X and Y are independent.
- ▶ Intuitively, X is amount of time till we see s events, and Y is amount of subsequent time till we see t more events.
- ▶ So $f_X(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{s-1}}{\Gamma(s)}$ and $f_Y(y) = \frac{\lambda e^{-\lambda y} (\lambda y)^{t-1}}{\Gamma(t)}$.
- ▶ Now $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy$.
- ▶ Up to an a -independent multiplicative constant, this is

$$\int_0^a e^{-\lambda(a-y)} (a-y)^{s-1} e^{-\lambda y} y^{t-1} dy = e^{-\lambda a} \int_0^a (a-y)^{s-1} y^{t-1} dy.$$

- ▶ Letting $x = y/a$, this becomes $e^{-\lambda a} a^{s+t-1} \int_0^1 (1-x)^{s-1} x^{t-1} dx$.

Summing independent gamma random variables

- ▶ Say X is gamma (λ, s) , Y is gamma (λ, t) , and X and Y are independent.
- ▶ Intuitively, X is amount of time till we see s events, and Y is amount of subsequent time till we see t more events.
- ▶ So $f_X(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{s-1}}{\Gamma(s)}$ and $f_Y(y) = \frac{\lambda e^{-\lambda y} (\lambda y)^{t-1}}{\Gamma(t)}$.
- ▶ Now $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy$.
- ▶ Up to an a -independent multiplicative constant, this is

$$\int_0^a e^{-\lambda(a-y)} (a-y)^{s-1} e^{-\lambda y} y^{t-1} dy = e^{-\lambda a} \int_0^a (a-y)^{s-1} y^{t-1} dy.$$

- ▶ Letting $x = y/a$, this becomes $e^{-\lambda a} a^{s+t-1} \int_0^1 (1-x)^{s-1} x^{t-1} dx$.
- ▶ This is (up to multiplicative constant) $e^{-\lambda a} a^{s+t-1}$. Constant must be such that integral from $-\infty$ to ∞ is 1. Conclude that $X + Y$ is gamma $(\lambda, s + t)$.

Summing two normal variables

- ▶ X is normal with mean zero, variance σ_1^2 , Y is normal with mean zero, variance σ_2^2 .

Summing two normal variables

- ▶ X is normal with mean zero, variance σ_1^2 , Y is normal with mean zero, variance σ_2^2 .
- ▶ $f_X(x) = \frac{1}{\sqrt{2\pi\sigma_1}} e^{\frac{-x^2}{2\sigma_1^2}}$ and $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_2}} e^{\frac{-y^2}{2\sigma_2^2}}$.

Summing two normal variables

- ▶ X is normal with mean zero, variance σ_1^2 , Y is normal with mean zero, variance σ_2^2 .
- ▶ $f_X(x) = \frac{1}{\sqrt{2\pi\sigma_1}} e^{\frac{-x^2}{2\sigma_1^2}}$ and $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_2}} e^{\frac{-y^2}{2\sigma_2^2}}$.
- ▶ We just need to compute $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$.

Summing two normal variables

- ▶ X is normal with mean zero, variance σ_1^2 , Y is normal with mean zero, variance σ_2^2 .
- ▶ $f_X(x) = \frac{1}{\sqrt{2\pi\sigma_1}} e^{\frac{-x^2}{2\sigma_1^2}}$ and $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_2}} e^{\frac{-y^2}{2\sigma_2^2}}$.
- ▶ We just need to compute $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$.
- ▶ We could compute this directly.

Summing two normal variables

- ▶ X is normal with mean zero, variance σ_1^2 , Y is normal with mean zero, variance σ_2^2 .
- ▶ $f_X(x) = \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{x^2}{2\sigma_1^2}}$ and $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_2}} e^{-\frac{y^2}{2\sigma_2^2}}$.
- ▶ We just need to compute $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$.
- ▶ We could compute this directly.
- ▶ Or we could argue with a multi-dimensional bell curve picture that if X and Y have variance 1 then $f_{\sigma_1 X + \sigma_2 Y}$ is the density of a normal random variable (and note that variances and expectations are additive).

Summing two normal variables

- ▶ X is normal with mean zero, variance σ_1^2 , Y is normal with mean zero, variance σ_2^2 .
- ▶ $f_X(x) = \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{x^2}{2\sigma_1^2}}$ and $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_2}} e^{-\frac{y^2}{2\sigma_2^2}}$.
- ▶ We just need to compute $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$.
- ▶ We could compute this directly.
- ▶ Or we could argue with a multi-dimensional bell curve picture that if X and Y have variance 1 then $f_{\sigma_1 X + \sigma_2 Y}$ is the density of a normal random variable (and note that variances and expectations are additive).
- ▶ Or use fact that if $A_i \in \{-1, 1\}$ are i.i.d. coin tosses then $\frac{1}{\sqrt{N}} \sum_{i=1}^N A_i$ is approximately normal with variance σ^2 when N is large.

Summing two normal variables

- ▶ X is normal with mean zero, variance σ_1^2 , Y is normal with mean zero, variance σ_2^2 .
- ▶ $f_X(x) = \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{x^2}{2\sigma_1^2}}$ and $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_2}} e^{-\frac{y^2}{2\sigma_2^2}}$.
- ▶ We just need to compute $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy$.
- ▶ We could compute this directly.
- ▶ Or we could argue with a multi-dimensional bell curve picture that if X and Y have variance 1 then $f_{\sigma_1 X + \sigma_2 Y}$ is the density of a normal random variable (and note that variances and expectations are additive).
- ▶ Or use fact that if $A_i \in \{-1, 1\}$ are i.i.d. coin tosses then $\frac{1}{\sqrt{N}} \sum_{i=1}^N A_i$ is approximately normal with variance σ^2 when N is large.
- ▶ Generally: if independent random variables X_j are normal (μ_j, σ_j^2) then $\sum_{j=1}^n X_j$ is normal $(\sum_{j=1}^n \mu_j, \sum_{j=1}^n \sigma_j^2)$.

- ▶ Sum of an independent binomial (m, p) and binomial (n, p) ?

- ▶ Sum of an independent binomial (m, p) and binomial (n, p) ?
- ▶ Yes, binomial $(m + n, p)$. Can be seen from coin toss interpretation.

- ▶ Sum of an independent binomial (m, p) and binomial (n, p) ?
- ▶ Yes, binomial $(m + n, p)$. Can be seen from coin toss interpretation.
- ▶ Sum of independent Poisson λ_1 and Poisson λ_2 ?

- ▶ Sum of an independent binomial (m, p) and binomial (n, p) ?
- ▶ Yes, binomial $(m + n, p)$. Can be seen from coin toss interpretation.
- ▶ Sum of independent Poisson λ_1 and Poisson λ_2 ?
- ▶ Yes, Poisson $\lambda_1 + \lambda_2$. Can be seen from Poisson point process interpretation.