

Sums and averages of iid random variables

- consider a collection of independent and identically distributed random variables from some common distribution P :

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} P.$$

- assume that the X_i have finite mean and variance. Since they are identically distributed, let $\mu = E(X_1)$ and $\sigma^2 = \text{var}(X_1)$ denote the common mean and variance that all of the X_i share.
- Define two new random variables:

$$\text{sum: } S_n = \sum_{i=1}^n X_i$$

$$\text{average: } \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} S_n.$$

- We seek to itemize the properties of these new random variables:

1. mean of S_n, \bar{X}_n ?

2. variance of S_n, \bar{X}_n ?

3. distribution of S_n, \bar{X}_n ?

4. what happens when $n \rightarrow \infty$?

1. What's the mean of the sum and the average?

Theorem (linearity of expectation)

- X_1, X_2, \dots, X_n are possibly dependent random variables with finite means: $E[|X_i|] < \infty \quad \forall i=1, \dots, n$;
- $a_1, a_2, \dots, a_n \in \mathbb{R}$ are constants;
- X_i could have any type, any distribution, any dependence.

Then

$$E\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i E(X_i).$$

Partial proof Consider the $n=2$ case with jointly absolutely continuous $(X, Y) \sim f_{XY}$, where f_{XY} is the joint pdf, and X, Y could be dependent. For any $a, b \in \mathbb{R}$, we want to show that $E[aX + bY] = aE(X) + bE(Y)$.

Recall multivariable LOTUS: $E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) dx dy$.

We apply this to the linear transformation $g(x, y) = ax + by$.

$$\begin{aligned} E[aX + bY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (ax + by) f_{XY}(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [ax f_{XY}(x, y) + by f_{XY}(x, y)] dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ax f_{XY}(x, y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} by f_{XY}(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ax f_{XY}(x, y) dy dx + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} by f_{XY}(x, y) dx dy \end{aligned}$$

(continued)

$$\begin{aligned} E[aX + bY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ax f_{XY}(x, y) dy dx + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} by f_{XY}(x, y) dx dy \\ &= a \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{XY}(x, y) dy dx + b \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy \\ &= a \int_{-\infty}^{\infty} x f_X(x) dx + b \int_{-\infty}^{\infty} y f_Y(y) dy \\ &= a E(X) + b E(Y). \end{aligned}$$

This covers the $n=2$ case. The general case follows by induction. \square

Corollaries . . .

$$E(S_n) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n \mu = \underbrace{\mu + \mu + \dots + \mu}_{n \text{ times}} = \boxed{n\mu}.$$

$$E(\bar{X}_n) = E\left(\frac{1}{n} S_n\right) = \frac{1}{n} E(S_n) = \frac{1}{n} n\mu = \boxed{\mu}.$$

So, when $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} P$ with common mean $\mu = E(X_i)$

then no matter what P is, we know

$$E\left(\sum_{i=1}^n X_i\right) = n\mu.$$

$$E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \mu.$$

2. What's the variance?

Theorem: If $a, b \in \mathbb{R}$ constant and X, Y are possibly dependent random variables of any type, then

$$\text{var}(aX + bY) = a^2 \text{var}(X) + b^2 \text{var}(Y) + 2ba \text{cov}(X, Y).$$

Proof:

$$\begin{aligned} \text{var}(aX + bY) &= E\left[\left((aX + bY) - E(aX + bY)\right)^2\right] \\ &= E\left[\left(aX + bY - aE(X) - bE(Y)\right)^2\right] \\ &= E\left[\left(a[X - E(X)] + b[Y - E(Y)]\right)^2\right] \\ &= E\left[a^2(X - E(X))^2 + b^2(Y - E(Y))^2 + 2ab(X - E(X))(Y - E(Y))\right] \\ &= a^2 E[(X - E(X))^2] + b^2 E[(Y - E(Y))^2] + 2ab E[(X - E(X))(Y - E(Y))] \\ &= a^2 \text{var}(X) + b^2 \text{var}(Y) + 2ab \text{cov}(X, Y). \end{aligned}$$

Corollary: If X, Y independent, then $\text{cov}(X, Y) = 0$, so

$$\text{var}(aX + bY) = a^2 \text{var}(X) + b^2 \text{var}(Y).$$

Corollary: If X_1, X_2, \dots, X_n independent, then by induction...

$$\text{var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{var}(X_i)$$

For us . . .

$$X_1, X_2, \dots, X_n \stackrel{iid}{\sim} P \quad \left(\Rightarrow \sigma^2 = \text{var}(X_1) \right)$$

$$S_n = \sum_{i=1}^n X_i$$

$$\bar{X}_n = \frac{1}{n} S_n$$

So

$$\text{var}(S_n) = \text{var}\left(\sum_{i=1}^n X_i\right)$$

$$= \sum_{i=1}^n \text{var}(X_i)$$

$$= \sum_{i=1}^n \sigma^2$$

$$= n \cdot \sigma^2$$

$$\text{var}(\bar{X}_n) = \text{var}\left(\frac{1}{n} S_n\right)$$

$$= \left(\frac{1}{n}\right)^2 \text{var}(S_n)$$

$$= \frac{1}{n^2} n \sigma^2$$

$$= \frac{\sigma^2}{n} .$$

true for
any P ,
as long as the
 X_i are independent
and the variance
is actually finite.

3. What is the entire distribution of S_n or \bar{X}_n ?

For the mean and the variance of S_n and \bar{X}_n , we got generic results that applied no matter what the underlying P was. We won't quite be able to do that for this third question. We'll just work some examples.

Recall: If $X \sim M_X$, $a, b \in \mathbb{R}$ constant, and $Y = aX + b$,
then $M_Y(t) = e^{bt} M_X(at)$. ← mgf


Proof:

$$\begin{aligned} M_Y(t) &= E[e^{tY}] = E[e^{t(aX+b)}] \\ &= E[e^{atX + bt}] \\ &= E[e^{bt} e^{atX}] \\ &= e^{bt} E[e^{atX}] \\ &= e^{bt} M_X(at). \end{aligned}$$

Theorem: If $X \sim M_X$ and $Y \sim M_Y$ are independent and $a, b \in \mathbb{R}$ are constant, then

$$\begin{aligned} M_{aX+bY}(t) &= E[e^{t(aX+bY)}] \\ &= E[e^{taX + btY}] \\ &= E[e^{atX} e^{btY}] \\ &= E[e^{atX}] E[e^{btY}] \\ &= M_X(at) M_Y(bt) \end{aligned}$$

expectation of a product is the product of the expectations by independence.



Corollary: X_1, X_2, \dots, X_n independent and a_1, a_2, \dots, a_n constant, then by induction...

$$\begin{aligned} M_{\sum_{i=1}^n a_i X_i}(t) &= \prod_{i=1}^n M_{X_i}(a_i t) \\ &= M_{X_1}(a_1 t) M_{X_2}(a_2 t) \cdots M_{X_n}(a_n t) \end{aligned}$$

So... $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} M$ ← common mgf shared by each of the X_i

$$M_{S_n}(t) = \prod_{i=1}^n M(t) = M(t)^n$$

$$M_{\bar{X}_n}(t) = M_{\frac{1}{n} S_n}(t) = M_{S_n}\left(\frac{t}{n}\right) = M\left(\frac{t}{n}\right)^n$$

Based on these, we can determine the distribution of the sum and the average in special cases.

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Poisson}(\lambda) \quad M(t) = e^{\lambda(e^t - 1)}$$

$$\left. \begin{aligned} M_{S_n}(t) &= M(t)^n \\ &= \left[e^{\lambda(e^t - 1)} \right]^n \\ &= e^{n\lambda(e^t - 1)} \end{aligned} \right\} S_n \sim \text{Poisson}(n\lambda)$$

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Gamma}(\alpha, \beta) \quad M(t) = \left(\frac{\beta}{\beta - t} \right)^\alpha$$

$$M_{S_n}(t) = M(t)^n = \left[\left(\frac{\beta}{\beta - t} \right)^\alpha \right]^n = \left(\frac{\beta}{\beta - t} \right)^{n\alpha}$$

$$\begin{aligned} M_{\bar{X}_n}(t) &= M(t/n)^n = \left[\left(\frac{\beta}{\beta - t/n} \right)^\alpha \right]^n \\ &= \left(\frac{\beta}{\frac{1}{n}(n\beta - t)} \right)^{n\alpha} \\ &= \left(\frac{n\beta}{n\beta - t} \right)^{n\alpha} \end{aligned}$$

So

$$S_n \sim \text{Gamma}(n\alpha, \beta)$$

$$\bar{X}_n \sim \text{Gamma}(n\alpha, n\beta)$$

$$X_1, X_2, \dots, X_n \stackrel{iid}{\sim} N(\mu, \sigma^2) \quad M(t) = e^{\mu t + \frac{\sigma^2}{2} t^2}$$

$$M_{S_n}(t) = M(t)^n = \left[e^{\mu t + \frac{\sigma^2}{2} t^2} \right]^n = e^{n\mu t + n\sigma^2 \frac{t^2}{2}}$$

$$\begin{aligned} M_{\bar{X}_n}(t) &= M(t/n)^n = \left[e^{\mu \frac{t}{n} + \frac{\sigma^2}{2} \left(\frac{t}{n}\right)^2} \right]^n \\ &= e^{n\mu \frac{t}{n} + n \frac{\sigma^2 t^2}{2 n^2}} \\ &= e^{\mu t + \frac{\sigma^2}{n} \frac{t^2}{2}} \end{aligned}$$

So...

$$S_n \sim N(n\mu, n\sigma^2)$$

$$\bar{X}_n \sim N(\mu, \sigma^2/n)$$

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Bern}(p)$$

$$\begin{aligned} M(t) &= E[e^{tX_1}] = (1-p)e^{t \cdot 0} + pe^{t \cdot 1} \\ &= (1-p) + pe^t \end{aligned}$$

And if $Y \sim \text{Binom}(n, p)$, then

$$M_Y(t) = E[e^{tY}] = \sum_{k=0}^n e^{tk} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^n \binom{n}{k} (e^t p)^k (1-p)^{n-k}$$

$$= \left[(1-p) + pe^t \right]^n$$

binomial
theorem
 $(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$

$$M_{S_n}(t) = M(t)^n = \left[(1-p) + pe^t \right]^n, \text{ so } \dots$$

$$S_n \sim \text{Binom}(n, p)$$

That's our first actual proof that the sum of iid Bernoullis is binomial.

Summary

Start with $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} P$ with common mean and variance $\mu = E(X_1)$ and $\sigma^2 = \text{var}(X_1)$.

Define new random variables

$$S_n = \sum_{i=1}^n X_i$$

$$\bar{X}_n = \frac{1}{n} S_n$$

means...

$$E(S_n) = n\mu$$

$$E(\bar{X}_n) = \mu$$

variances...

$$\text{var}(S_n) = n\sigma^2$$

$$\text{var}(\bar{X}_n) = \sigma^2/n$$

distribution

$$P = \text{Poisson}(\lambda) \implies S_n \sim \text{Poisson}(n\lambda)$$

$$P = \text{Bernoulli}(p) \implies S_n \sim \text{Binom}(n, p)$$

$$P = \text{Gamma}(\alpha, \beta) \implies S_n \sim \text{Gamma}(n\alpha, \beta)$$

$$\bar{X}_n \sim \text{Gamma}(n\alpha, n\beta)$$

$$P = N(\mu, \sigma^2) \implies S_n \sim N(n\mu, n\sigma^2)$$

$$\bar{X}_n \sim N(\mu, \sigma^2/n)$$