

Recap

- (Ch 4) Random variables
 - Expected value, variance and covariance
 - Towards the weak law of large numbers

Today

- (Ch 4) Random variables
 - Weak law of large numbers
 - Simulation examples
 - Continuous random variables

Towards the weak law of large numbers

- The weak law says that if we repeat an experiment many times, the average of the observations will “converge” to the expected value
- The weak law justifies using simulations (instead of calculations) to estimate the expected values of random variables

Markov's inequality

- For any random variable X and constant $a > 0$

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

- In words, a random variable is unlikely to have an absolute value much larger than the mean of its absolute value
- For example, if $a = 10E[|X|]$

$$P(|X| \geq 10E[|X|]) \leq 0.1$$

Chebyshev's inequality

- For any random variable X and constant $a > 0$

$$P(|X - E[X]| \geq a) \leq \frac{\text{var}[X]}{a^2}$$

- To rephrase, let $a = k\sigma$ where $\sigma = \text{std}[X]$

$$P(|X - E[X]| \geq k\sigma) \leq \frac{1}{k^2}$$

- In words, the probability that X is greater than k standard deviations from the mean is small

Proof of Chebyshev's inequality

- Apply Markov's inequality to $U = (X - E[X])^2$, so $E[U] = \text{var}[X]$

$$P(|U| \geq w) \leq \frac{E[|U|]}{w} = \frac{E[U]}{w} = \frac{\text{var}[X]}{w}$$

- Substitute $U = (X - E[X])^2$ and $w = a^2$

$$P\left(\underbrace{(X - E[X])^2}_{\geq a^2}\right) \leq \frac{\text{var}[X]}{a^2}$$

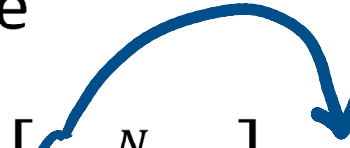
$$P(|X - E[X]| \geq a)$$

IID samples and the sample mean $E[x] = \sum_x x P(x)$

- Say you have a random variable X with probability distribution $P(X)$
- Say you **randomly** generate independent samples $\{x_i\}$ so that their histogram resembles $P(X)$ more closely as their number increases
- We call $\{x_i\}$ independent identically distributed **(IID) samples** of $P(X)$
- The **sample mean** of $\{x_i\}$ is a random variable: $X_N = \frac{1}{N} \sum_{i=1}^N x_i$

Expected value of sample mean

- By linearity of expected value

$$E[X_N] = E\left[\frac{1}{N}\sum_{i=1}^N x_i\right] = \frac{1}{N}\sum_{i=1}^N E[x_i]$$


- Since each x_i is a sample drawn from $P(X)$, we have $E[x_i] = E[X]$

$$E[X_N] = \frac{1}{N}\sum_{i=1}^N E[X] = E[X]$$

Variance of sample mean

- By the scaling property of variance and independence of samples x_i

$$\text{var}[X_N] = \text{var}\left[\frac{1}{N} \sum_{i=1}^N x_i\right] = \frac{1}{N^2} \text{var}\left[\sum_{i=1}^N x_i\right] = \frac{1}{N^2} \sum_{i=1}^N \text{var}[x_i]$$

- Since each x_i is drawn from $P(X)$, we have $\text{var}[x_i] = \text{var}[X]$

$$\text{var}[X_N] = \frac{1}{N^2} \sum_{i=1}^N \text{var}[X] = \frac{\text{var}[X]}{N}$$

Weak law of large numbers (WLLN)

- Given a random variable X with finite variance, probability distribution $P(X)$ and sample mean X_N
- For any positive number ϵ

$$\lim_{N \rightarrow \infty} P(|X_N - E[X]| \geq \epsilon) = 0$$

- In words: for a large enough set of IID samples, the sample mean X_N will be very close to the expected value $E[X]$ with high probability

Proof of WLLN

- Apply Chebyshev's inequality to the sample mean X_N

$$P(|X_N - E[X_N]| \geq \epsilon) \leq \frac{\text{var}[X_N]}{\epsilon^2}$$

- Substitute $E[X_N] = E[X]$ and substitute $\text{var}[X_N] = \text{var}[X]/N$

$$P(|X_N - E[X]| \geq \epsilon) \leq \frac{\text{var}[X]}{N\epsilon^2} \xrightarrow{N \rightarrow \infty} 0$$

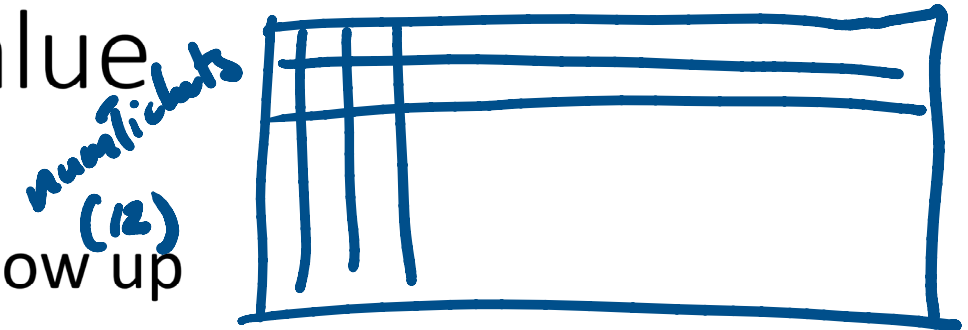
Simulation: airline overbooking example

An airline has a flight with 6 seats. They always sell 12 tickets for this flight and ticket holders show up independently with probability p . Plot the following quantities as a function of p .

- Expected number of ticket holders that show up
- Probability that the flight is overbooked
- Expected number of ticket holders who show up but don't fly given that the flight is overbooked

Approximating expected value

Expected number of ticket holders that show up



```
results = np.zeros((10, 2))
numTrials = 100000
numTickets = 12
numSeats = 6
for i,p in enumerate(np.linspace(0.1, 1.0, num=10)):
    arrivals = np.random.random((numTickets, numTrials)) < p
    numArrivals = arrivals.sum(axis=0)
    results[i] = [p, numArrivals.mean()]
results
```

[0.1, 0.2, 0.3, ..., 1.0]

Approximating probability

Probability that the flight is overbooked

```
results = np.zeros((10, 2))
numTrials = 100000
numTickets = 12
numSeats = 6
for i,p in enumerate(np.linspace(0.1, 1.0, num=10)):
    arrivals = np.random.random((numTickets, numTrials)) < p
    numArrivals = arrivals.sum(axis=0)
    indicatorOverbooked = numArrivals > numSeats
    results[i] = [p, indicatorOverbooked.mean()]
results
```

Approximating conditional expected value

Expected number of ticket holders who show up but don't fly given that the flight is overbooked

```
results = np.zeros((10, 2))
numTrials = 100000
numTickets = 12
numSeats = 6
for i,p in enumerate(np.linspace(0.1, 1.0, num=10)):
    arrivals = np.random.random((numTickets, numTrials)) < p
    numArrivals = arrivals.sum(axis=0)
    indicatorOverbooked = numArrivals > numSeats
    numDontFly = numArrivals[indicatorOverbooked] - numSeats
    results[i] = [p, numDontFly.mean()]
results
```