MM 225: AI and Data Science

Moments, convolution and law of large numbers

M P Gururajan and Hina A Gokhale

September 5, 2023

Lecture 9: Moments, convolution, and law of large numbers

Outline

Queues

- Suppose customers join a queue in such a way that the arrival is exponentially distributed with parameter λ
- Expected value of time between arrivals: $\frac{1}{\lambda}$
- Arrival rate: λ
- Service time: time customer has to wait before leaving
- Suppose service time is also exponentially distributed: with parameter μ
- Expected value of service time: $\frac{1}{\mu}$
- Service rate: μ
- Traffic intensity $\rho =$(arrival rate)$\times$(service time)$= \frac{\lambda}{\mu}$
- Traffic intensity: ratio of service time to arrival time
- $\rho > 1$: indefinitely large; $\rho < 1$: reasonable; $\rho = 1$: large queues but at times empty

Case of $\rho < 1$

- Suppose traffic intensity is less than 1
- Length of queue Z is a random variable with finite expected value $E(Z) = N$
- Time spent by each cutomer W is a random variable with finite expected value $E(W) = T$
- Customer expects N people ahead while joining the queue and λT behind while leaving
- At equilibrium these are the same: hence $N = \lambda T$
- Little's law for queues
- Queue length: geometric distribution
- Simulation of queue lengths and geometric distribution: python tutorial

Moments

Moments

- Consider our definition of mean and variance
- μ_k = k-th moment of X
- $\mu_k = E(X^k) = \sum_{j=1}^{\infty}(x_j)^k p(x_j)$ (provided the sum converges)
- Third and fourth moments: related to skewness and kurtosis
- Recall python tutorial where you have calculated the first two moments as well as skewness and kurtosis
- Remark: knowledge of all moments of X determines its distribution function completely
- Moment generating functions: check the textbook

Moments

- Concept of moments for continuous random variables
- $\mu_k = k$-th moment of X
- $\mu_k = E(X^k) = \int_{-\infty}^{\infty} (x)^k f_X(x) dx$ (provided the integral$\mu_k = E(X^k) = \int_{-\infty}^{\infty} (|x|)^k f_X(x) dx$ is finite)
- Third and fourth moments: related to skewness and kurtosis
- Recall python tutorial where you have calcualated the first two moments as well as skewness and kurtosis

Convolution

Convolution

- X and Y: independent discrete random variabls with distribution functions $m_1(x)$ and $m_2(x)$
- Let $Z = X + Y$
- What is the distribution function $m_3(x)$ of Z?
- What is the probability that Z takes the value z?
- Let X take a value k
- Z will take a value z only if Y takes the value $z - k$
- Event $Z = z$ is an union of pairwise disjoint events $X = k$ and $Y = z - k$
- This implies $P(Z = z) = \sum_{k=-\infty}^{k=\infty} P(X = k) \cdot P(Y = z - k)$

Definition

Definition

Let X and Y be two independent integer-valued random variables, with distribution functions $m_1(x)$ and $m_2(x)$ respectively. Then the convolution of $m_1(x)$ and $m_2(x)$ is the distribution function $m_3 = m_1 \star m_2$ given by $m_3(j) = \sum_k m_1(k) \cdot m_2(j-k)$, for $j = \cdots, -2, -1, 0, 1, 2, \cdots$. The function $m_3(x)$ is the distribution function of the random variable $Z = X + Y$.

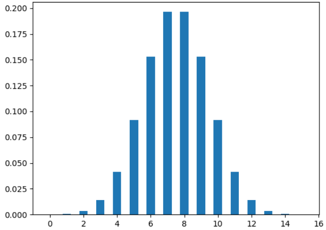
Remark

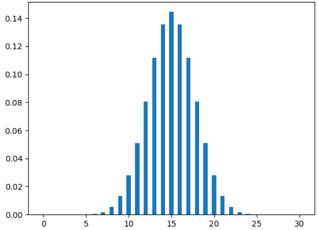
Convolution is commutative and associative

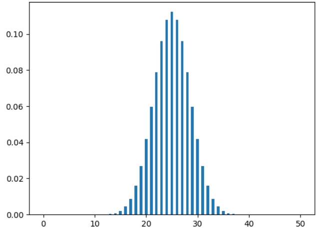
Summing random variables: Bernoulli trials

- Consider a Bernoulli trial; let ± 1 be the sample space (success and failure)
- The distribution function: $m(1) = p$; $m(-1) = q$
- Let $S_n = X_1 + X + 2 + \cdots + X_n$ where X are independent random variables with the same density function m
- What is the density function for S_n?
- $S_n = S_{n-1} + X_n$
- The distribution function for X_1 is m; hence, by induction, we can find the density function of S_n
- Simulation: home work!
- Bernoulli trials: binomial distribution
- $\binom{n}{k} p^k q^{n-k}$

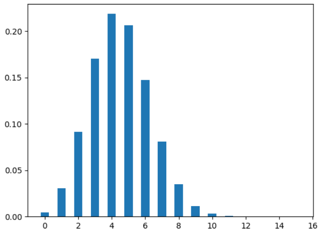
$p = 0.5,\ n = 15, 30, 50$

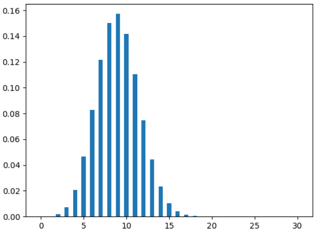

0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0 2 4 6 8 10 12 14 16

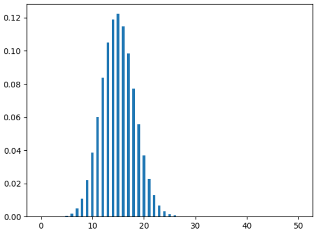

0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0 5 10 15 20 25 30


0.10
0.08
0.06
0.04
0.02
0.00
0 10 20 30 40 50

$p = 0.3,\ n = 15, 30, 50$


0.00
0.05
0.10
0.15
0.20
0
2
4
6
8
10
12
14
16


0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0.16
0
5
10
15
20
25
30


0.00
0.02
0.04
0.06
0.08
0.10
0.12
0
10
20
30
40
50

Summing random variables: Throw of dies

- Consider the throw of a fair die; let $(1, 2, 3, 4, 5, 6)$ be the sample space
- The distribution function: $m(1) = m(2) = m(3) = m(4) = m(5) = m(6) = \frac{1}{6}$;
- Let $S_n = X_1 + X + 2 + \cdots + X_n$ where X are independent random variables with the same density function m
- What is the density function for S_n?
- $S_n = S_{n-1} + X_n$
- The distribution function for X_1 is m; hence, by induction, we can find the density function of S_n
- Simulation: home work!

Strategy

- Recall binomial: for Bernoulli trials
- Die throw: how do we calculate the distribution function without simulation?
- Write a script to return the convolution r of distributions p and q
- Keep calling the function using the induction formula
- Do for $N = 10, 20, 30$
- Generate the distribution functions
- A problem for one of the python tutorial sessions!

Convolution of continuous random variables

Definition

Let X and Y be two continuous random variables with density functions $f(x)$ and $g(y)$, respectively. Assuming both $f(x)$ and $g(y)$ are defined for all real numbers, the convolution $f \star g$ of f and g is the function given by
$(f \star g)(z) = \int_{-\infty}^{\infty} f(z-y)g(y)dy = \int_{-\infty}^{\infty} g(z-x)f(x)dx.$

Theorem

Let X and Y be two independent random variables with density functions $f_X(x)$ and $f_Y(y)$ defined for all x. Then the sum $Z = X + Y$ is a random variable with density function $f_Z(z)$, where f_Z is the convolution of f_X and f_Y.

Convolution of two uniform random numbers

- $X + Y = Z$ where X and Y are uniform random variables
- $f_X(x) = f_Y(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$
- $f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy$
- Since $f_Y(y) = 1$ if $0 \le y \le 1$ and zero, otherwise, $f_Z(z) = \int_0^1 f_X(z-y) dy$
- Integrand is zero unless $0 \le z - y \le 1$;
- Or, $z - 1 \le y \le z$
- If $0 \le z \le 1$: $f_Z(z) = \int_0^z dy = z$
- If $1 < z \le 2$: $f_Z(z) = \int_{z-1}^1 dy = 2 - z$
- Hence, $f_Z(z) = \begin{cases} z & \text{if } 0 \le z \le 1 \\ 2 - z & \text{if } 1 \le z \le 2 \\ 0 & \text{otherwise} \end{cases}$

Convolution of two exponential random numbers

- $Z = X + Y$
- $f_X(x) = f_Y(x) = \begin{cases} \int_0^\infty \lambda \exp(-\lambda x) & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$
- If $z > 0$ $f_Z(z) = \int_{-\infty} \infty f_X(z-y) f_Y(y) dy = \int_0^z \lambda \exp(-\lambda(z-y))\lambda \exp(-\lambda y) dy$
- $f_Z(z) = \int_0^z \lambda^2 \exp(-\lambda z) dy = \lambda^2 z \exp(-\lambda z)$ if $z \geq 0$
- $f_Z(z) = 0$ if $z < 0$

Convolution of two normal random numbers

- Convolution of two independent normal random variables with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 is a normal random variable with mean $\mu = \mu_1 + \mu_2$ and variance $\sigma^2 = \sigma_1^2 + \sigma_2^2$
- Homework: prove this result for standard normal distribution!

Chi-squared distribution

- Sum of squared of n independent normally distributed random variables with mean zero and standard deviation unity: gamma density with $\lambda = \frac{1}{2}$ and $\beta = \frac{n}{2}$
- Chi-squared density with n degrees of freedom
- Use of Chi-quared density: for hypothesis testing and for comparing experimental data with theoretical distribution
- Python tutorial session: we will do more examples

Law of large numbers

Question

- If a random variable is distributed as normal, 95% of the data lies within 2σ where σ is the standard deviation.
- Yes / No

Speculation

- Suppose we do not know the distribution of a random variable. We can safely assume that 95% of the data will lie within 2 times standard deviation.

Answer

- **Link:** https://www.menti.com/alxoochtjz7x

- **Code:** Go to menti.com and use code 4183 4566

Results

Results of the poll

Markov's inequality

Theorem

If X is a random variable that takes non-negative values, then, for any $a > 0$,
$P(X \geq a) \leq \frac{E(X)}{a}$

Chebyeshev inequality

Theorem

If X is a discrete random variable with $E(X) = \mu$, then, for any $\epsilon > 0$, $P(|X - \mu| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2}$

Theorem

Let X be a continuous random variable with density function $f(x)$. Suppose X has a finite expected value $\mu = E(X)$ and finite variance $\sigma^2 = V(X)$. Then for any positive number $\epsilon > 0$ we have $P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$

Law of large numbers

Theorem

Let $X_1, X_2, \cdots, X_n$ be an independent trials process with finite expected value $\mu = E(X_j)$ and finite variance $\sigma^2 = V(X_j)$. Let $S_n = X_1 + X_2 + \cdots + X_n$. Then for any $\epsilon > 0$,
$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \to 0$ as $n \to \infty$. Or, $P\left(\left|\frac{S_n}{n} - \mu\right| \leq \epsilon\right) \to 1$

Theorem

Let $X_1, X_2, \cdots, X_n$ be an independent trials process with a continuous density function f, finite expected value μ and finite variance σ^2. Let $S_n = X_1 + X_2 + \cdots + X_n$. Then for any $\epsilon > 0$, $\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0$. Or, $\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \leq \epsilon\right) = 1$

THANK YOU!!!