INEQUALITIES, LLSE 11

COMPUTER SCIENCE MENTORS 70

April 16-18, 2018

1 Normal Distribution

Gaussian (Normal) Distribution: $N(\mu, \sigma^2)$

- Mean: μ
- Variance: σ^2
- $f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$
- If $X \sim N(\mu, \sigma)$ then $b + X \sim N(b + \mu, \sigma^2)$
- If $X \sim N(\mu, \sigma)$ then $cX \sim N(c\mu, c^2\sigma^2)$
- If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ and X and Y are independent then

$$X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

- The CLT (Central Limit Theorem) states that for a sequence of iid random variables: $X_1, X_2, ..., X_n$, each with mean μ and variance σ^2,

$$\frac{X_1 + X_2 + ... + X_n - n\mu}{\sigma\sqrt{n}}$$

approaches the standard normal distribution $Z \sim N(0, 1)$

1.1 Questions

1. **Coin Flipping**
Suppose you have two coins, one that has heads on both sides and another that has tails on both sides. You pick one of the two coins uniformly at random and flip it. You repeat this process 400 times, each time picking one of the two coins uniformly at

random and then flipping it, for a total of 400 flips.
Use the CLT to approximate the probability of getting more than 220 heads.

2 More Continuous Practice

2.1 Questions

1. Alice and Bob are throwing baseballs and they want to see who can throw a baseball further. The distance Alice throws a baseball is modeled as a uniform distribution between 0 and 5 while the distance Bob throws a baseball is modeled as a uniform distribution between 0 and 10. Assume Alice and Bob throw independently, what is the probability that Bob's throw will be further than Alice's?
2. Now Alice has improved her throwing abilities and her throwing distance is now also uniform on the interval 0 to 10, which is the same as Bob.
   Given that Alice's throw was greater than 5, what is the probability that Bob throws further than her?

3 Inequalities

3.1 Introduction

Markov's Inequality
For a non-negative random variable X with expectation $\mathrm{E}(X) = \mu$, and any $\alpha > 0$:

$$\mathrm{P}[X \geq \alpha] \leq \frac{\mathrm{E}(X)}{\alpha}$$

Proof of Markov's Inequality

$$\begin{aligned} E(X) &= \sum_a a * Pr[X = a] \\ &\geq \sum_{a \geq \alpha} a * Pr[X = a] \\ &\geq \alpha \sum_{a \geq \alpha} Pr[X = a] \\ &= \alpha Pr[X \geq a] \end{aligned}$$

Alternate Proof of Markov's Inequality
Consider the indicator random variable Y which equals 1 if $X \geq a$ and 0 otherwise. Now consider the relationship between X and aY

- If $X < a$, then $Y = 0$, which means $aY = a * 0 = 0$
  Because X is a non-negative random variable, $X \geq 0$, so $aY \leq X$ in this case.
- If $X \geq a$, then $Y = 1$, which means $aY = a * 1 = a \leq X$

Thus, we have $aY \leq X$.
We take expectation on both sides to get:

$$E[aY] \leq X$$

$$aE[Y] \leq E[X]$$

$$E[Y] \leq \frac{E[X]}{a}$$

Now we note that the expectation of an indicator random variable is the probability that it is equal to 1 and we have the proof:

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

Chebyshev's Inequality
For a random variable X with expectation $\mathrm{E}(X) = \mu$, and any $\alpha > 0$:

$$\mathrm{P}[|X - \mu| \geq \alpha] \leq \frac{\mathrm{Var}(X)}{\alpha^2}$$

Chebyshev's Inequality can be used to estimate the mean of an unknown distribution. Often times we do not know the true mean, so we take many samples $X_1, X_2, ..., X_n$. Our sample mean is thus $S_n = \frac{X_1+X_2+...+X_n}{n}$
We can upper-bound the probability of that our sample mean deviates too much from our true mean as:

$$P(|\hat{\mu} - \mu| > \epsilon) \leq \delta$$

where ϵ is known as our error and δ is known as our confidence.

3.2 Questions

1. Let X be the sum of 20 i.i.d. Poisson random variables $X_1, \ldots, X_{20}$ with $\mathrm{E}(X_i) = 1$. Find an upper bound of $\mathrm{P}[X \geq 26]$ using,

   (a) Markov's inequality:

   (b) Chebyshev's inequality:

2. **Bound It**
   A random variable X is always strictly larger than -100. You know that $\mathrm{E}(X) = -60$. Give the best upper bound you can on $\mathrm{P}[X \geq -20]$.

3. The citizens of the country USD (the United States of Drumpf) vote in the following manner for their presidential election: if the country is liberal, then each citizen votes for a liberal candidate with probability p and a conservative candidate with probability $1-p$, while if the country is conservative, then each citizen votes for a conservative candidate with probability p and a liberal candidate with probability $1-p$. After the election, the country is declared to be of the party with the majority of the votes.

   (a) Assume that $p = \frac{3}{4}$ and suppose that 100 citizens of USD vote in the election and that USD is known to be conservative. Provide a tight bound on the probability that it is declared to be a Liberal country.

   (b) Now let p be unknown; we wish to estimate it. Using Chebyshev's Inequality, determine the number of voters necessary to determine p within an error of 0.01, with probability at least 0.95.

4. **Squirrel Standard Deviation**
   As we all know, Berkeley squirrels are extremely fat and cute. The average squirrel is 40% body fat. The standard deviation of body fat is 5%. Provide an upper bound on the probability that a randomly trapped squirrel is either too skinny or too fat? A skinny squirrel has less than 27.5% body fat, and a fat squirrel has more than 52.5% body fat?

5. Consider a random variable Y with expectation μ whose maximum value is $\frac{3\mu}{2}$, prove that the probability that Y is 0 is at most $\frac{1}{3}$.

6. Let $X_1, X_2, ..., X_n$ be n iid Geometric random variables with parameter p. Using Chebyshev's inequality, provide an upper-bound on: $P(|\frac{X_1+X_2+...+X_n}{n} - \frac{1}{p})| \geq a)$
   Recall the variance for a Geometric Distribution with parameter p is $\frac{1-p}{p^2}$

7. Suppose we have a sequence of iid random variables $X_1, X_2, ..., X_n$
   Let $A_n = \frac{X_1+X_2+...+X_n}{n}$ be the sample mean.
   Show that the true mean of $X_i = \mu$ is within the interval $[\mu - 4.5\frac{\sigma}{\sqrt{n}}, \mu + 4.5\frac{\sigma}{\sqrt{n}}]$ with 95% probability.

4 Covariance

4.1 Introduction

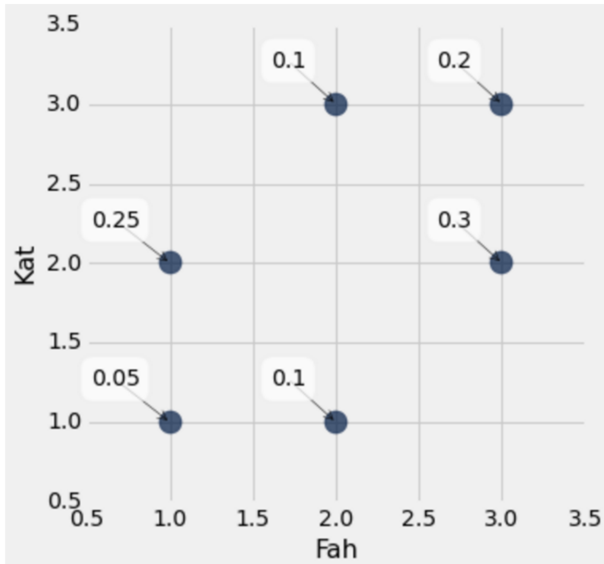
The **covariance** of two random variables X and Y is defined as:

$$\text{Cov}(X, Y) := \text{E}((X - \text{E}(X)) \cdot (Y - \text{E}(Y)))$$

4.2 Questions

1. Prove that $\text{Cov}(X, X) = \text{Var}(X)$:

2. Consider the following distribution with random variables Fah and Kat:



   Find the covariance of Fah and Kat.

3. Prove that if X and Y are independent, then $\text{Cov}(X, Y) = 0$:

4. Prove that the converse is not necessarily true. In other words, give an example of 2 random variables whose covariance is 0 but are not independent.

5. Roll 2 dice. Let A be the number of 6's you get, and B be the number of 5's, find $\text{Cov}(A, B)$

6. Prove that $\text{Cov}(X+Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$:

5 Linear Least Squares Estimator

Theorem: Consider two random variables, X, Y with a given distribution $\text{P}[X = x, Y = y]$. Then

$$\text{L}[Y|X] = \text{E}(Y) + \frac{\text{Cov}(X, Y)}{\text{Var}(X)}(X - \text{E}(X))$$

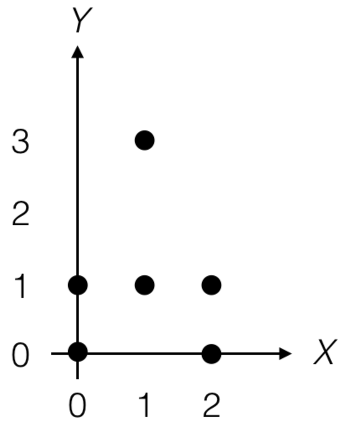
5.1 Questions

1. Assume that

$$Y = \alpha X + Z$$

where X and Z are independent and $\text{E}(X) = \text{E}(Z) = 0$. Find $\text{L}[X|Y]$.

2. The figure below shows the six equally likely values of the random pair (X, Y). Specify the functions of:
   - $L[Y \mid X]$
   - $E(X \mid Y)$
   - $L[X \mid Y]$
   - $E(Y \mid X)$


Y
3
2
1
0
X
0 1 2