INEQUALITIES, LLSE 11

COMPUTER SCIENCE MENTORS 70

April 16-18, 2018

1 Inequalities

1.1 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 $\alpha > 0$:

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

1.2 Questions

1. Use Markov's to prove Chebyshev's Inequality:

2. 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:

3. **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 $\mathrm{P}[X \geq -20]$.

4. 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.

5. **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?

6. 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 $\frac{1}{3}$.

7. 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 $\frac{1-p}{p^2}$

8. 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.

1.3 Covariance

1.4 Introduction

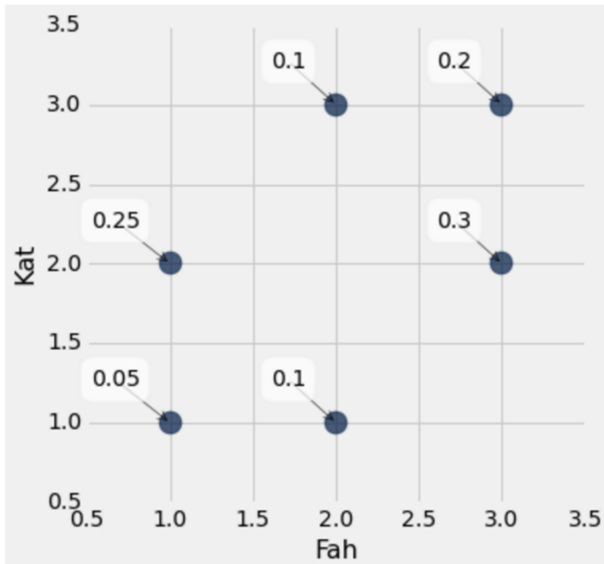
The **covariance** of two random variables X and Y is defined as:

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

1.5 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 $\text{Cov}(X+Y,Z) = \text{Cov}(X,Z) + \text{Cov}(Y,Z)$:

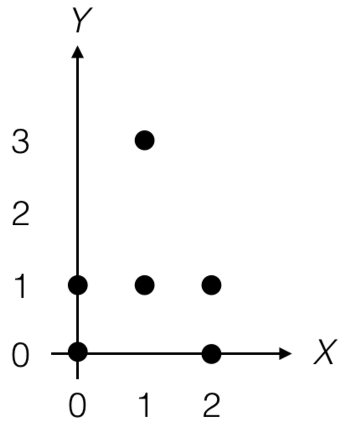
2 Linear Least Squares Estimator

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

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

2.1 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