Today: ASV 9.1-9.2

Next: ASV 9.3 (the end)

This week:

- homework 8 due Friday, December 6, 11:59 pm

- midterm 2 regrades: Tuesday, Dec 3, 8am-11pm
Example (Exercise 8.54)

Let $X, Y$ be r.v.'s satisfying

$$
E(X) = 2, E(Y) = 1, E(X^2) = 5, E(Y^2) = 10, E(XY) = 1
$$

(a) Compute $\text{Corr}(X, Y)$

$$
\text{Var}(X) = E(X^2) - (E(X))^2 = 5 - 2^2 = 1
\text{Var}(Y) = 9
\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 1 - 2 \cdot 1 = -1
$$

$$
\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{-1}{\sqrt{1 \cdot 9}} = -\frac{1}{3}
$$

(b) Find $c \in \mathbb{R}$ such that $X$ and $X + cy$ are uncorrelated

$$
\text{Cov}(X, X + cy) = \text{Cov}(X, X) + c \text{Cov}(X, Y) = 1 + c(-1) = 1 - c = 0
\Rightarrow c = 1
$$
**Interesting fact: expectation of nonnegative r.v.'s**

Let $X$ be a cont. r.v., $X \geq 0$. Then

$$E(X) = \int_{0}^{\infty} x f_X(x) dx$$

where $f_X(x)$ is the probability density function of $X$.

If $X$ is discrete and non-negative,

$$E(X) = \sum_{k=0}^{\infty} P(X \geq k)$$
Estimating tail probabilities: Markov's inequality

Thm. (Monotonicity of expectation).

(a) Let $Z$ be a r.v. $P(Z > 0) = 1$. Then $E(Z) \geq 0$

(b) Let $X, Y$ be r.v.'s, $P(X \geq Y) = 1$. Then $E(X) \geq E(Y)$

Proof. Apply (a) to r.v. $X - Y$

Markov's inequality

Thm. Let $X$ be a r.v., $P(X \geq 0) = 1$ ($X \geq 0$ almost surely).

Then for any $c > 0$ $P(X \geq c) \leq \frac{E(X)}{c}$

Proof. $X = X \cdot 1 \geq X(\omega) \cdot \mathbb{1}_{\{X \geq c\}}(\omega) \geq c \cdot \mathbb{1}_{\{X \geq c\}} \Rightarrow X \geq c \cdot \mathbb{1}_{\{X \geq c\}}$

$\Rightarrow E(X) \geq E(c \cdot \mathbb{1}_{\{X \geq c\}}) = c \cdot E(\mathbb{1}_{\{X \geq c\}}) = c \cdot P(X \geq c)$

Alternatively, $E(X) = \int_0^\infty P(X \geq s) ds$
Estimating tail probabilities: Chebyshev's inequality

**Thm.** Let $X$ be a r.v., $E(X) = \mu$, $\text{Var}(X) = \sigma^2$. Then for any $c > 0$

$$P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

**Proof.** $P(|X-\mu| \geq c) = P((X-\mu)^2 \geq c^2)$

$(X-\mu)^2 \geq 0 \implies \text{apply Markov's inequality}$

$$P((X-\mu)^2 \geq c^2) \leq \frac{E((X-\mu)^2)}{c^2} = \frac{\text{Var}(X)}{c^2} = \frac{\sigma^2}{c^2}$$

**Immediate corollaries:**

$$P(X-\mu \geq c) \leq P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

$$P(X-\mu \leq -c) \leq P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}$$
Examples

(Example 9.3) Markov's inequality: sometimes useful, not always
Let $X \sim \text{Ber}(p)$. $P(X \geq 1) = P(X \geq 0.01) = P(X = 1) = p$.

Estimate these probabilities using Markov's inequality

\[
P(X \geq 1) \leq \frac{E(X)}{1} = p, \quad P(X \geq 0.01) \leq \frac{E(X)}{0.01} = 100p
\]

(Exercise 9.2). $X \sim \text{Exp}(\frac{1}{2})$. $P(X \geq 6)$?

$E(X) = 2, \ Var(X) = 4$

(a) Markov: $P(X \geq 6) = \frac{E(X)}{6} = \frac{2}{6} = \frac{1}{3}$

(b) Chebyshev: $P(X \geq 6) = P(X - 2 \geq 4) \leq \frac{\text{Var}(X)}{16} = \frac{4}{16} = \frac{1}{4}$

(c) exact value: $P(X \geq 6) = e^{-\frac{1}{2}} \cdot e^{-3} = e^{-3} = 0.05$
(Exercise 9.9) \( X_i \) = amount of money earned by a food truck on day \( i \). From past experience \( E(X_i) = 5000 \)

(a) Estimate \( P(X_i \geq 7000) \).

Markov: \( P(X_i \geq 7000) \leq \frac{E(X_i)}{7000} = \frac{5000}{7000} = \frac{5}{7} \)

(b) Estimate \( P(X_i \geq 7000) \) if we know \( \text{Var}(X_i) = 4500 \)

Chebyshev: \( P(X_i \geq 7000) = P(X_i - 5000 \geq 2000) = \frac{4500}{(2000)^2} = \frac{9}{8000} \approx 0.001 \)

(c) Assume that \( X_i \)'s are independent. Let \( \overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \)

How big \( n \) should be to have \( P(4950 \leq \overline{X}_n \leq 5050) \geq 0.95 \)?

\[
E(\overline{X}_n) = E(X_i) = 5000, \quad \text{Var}(\overline{X}_n) = \frac{\sum_i \text{Var}(X_i)}{n^2} = \frac{4500}{n}
\]

\[
P(4950 \leq \overline{X}_n \leq 5050) = P(\overline{X}_n - 5000 \leq 250) \geq 1 - \frac{\text{Var}(\overline{X}_n)}{250^2} = 1 - \frac{4500}{25000} \geq 0.95 \implies n \geq 36
\]