Theory

Sunday, November 3, 2024 3:26 PM

- **Markov Inequality**

Let X r.v. such that $X \geq 0$. Then, $\forall t > 0$

$$P(X \geq t) \leq \frac{EX}{t}$$

- Upper bound of the above probability using the expected value
- Lets note the expected value as $\mu = EX$. Then, if t is a multiple of the expected value, i.e. $t = k\mu$, then:

$$P(X \geq k\mu) \leq \frac{\mu}{k\mu} \Rightarrow \boxed{P(X \geq k\mu) \leq \frac{1}{k}}$$

- If $k=1 \Rightarrow P(X \geq \mu) \leq 1$ (trivial)
- If $k=2 \Rightarrow P(X \geq 2\mu) \leq \frac{1}{2}$
- If $k=3 \Rightarrow P(X \geq 3\mu) \leq \frac{1}{3}$

...

- **Variance**

- Expected value → Central Tendency Measure (shows the value which the values of the distribution tend to get)
- Variance → Variability measure (shows how much the values diverge or the spread of the distribution)

Ορισμός: $\mathrm{Var}X = E[X - EX]^2$

divergence of the r.v. from the expected value

Thus, the expected value of the divergence of the r.v. from the expected value

- Properties

1) $\mathrm{Var}X = EX^2 - (EX)^2$
2) If X, Y independent $\Rightarrow \mathrm{Var}(X+Y) = \mathrm{Var}X + \mathrm{Var}Y$
3) $\mathrm{Var}[aX+b] = a^2 \mathrm{Var}X$

- **Standard Deviation**

Usually, variance may not be that meaningful, as the measurement units of variance are squared.

Definition: $\sigma = \sqrt{\mathrm{Var}X}$ (now this has the same units of measurement)

- **Covariance**

Definition: $Cov(X,Y) = E[XY] - E[X]\cdot E[Y]$

Usage: If X, Y independent $E(XY) = E(X)\cdot E(Y) \Rightarrow \boxed{Cov(X,Y) = 0}$

Note: Can be used to estimate if it is possible that X and Y are independent. However, $Cov(XY)=0$ doesn't guarantee independency. I.e., the opposite doesn't always hold

Properties

1) $Cov(X,Y) = Cov(Y,X)$
2) $Cov(X,X) = Var X$
3) $Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + \sum_{i}\sum_{j\neq i} Cov(X_i, X_j)$

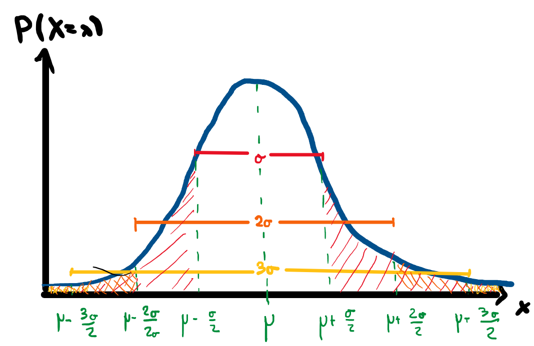
- **Chebyshev Inequality**

Definition: $P(|X-\mu| \geq t) \leq \frac{Var X}{t^2}$

Usage:

- Αν $Var X \downarrow \Rightarrow P(|X-\mu| \geq t) \downarrow$
- Αν $t \downarrow \Rightarrow \frac{Var X}{t^2} \uparrow$

(the smaller the variance, the larger the concentration around the expected value)



- Αν $t = \sigma \Rightarrow P(|X-\mu| \geq \sigma) \leq \frac{\sigma^2}{\sigma^2} = 1$
- Αν $t = 2\sigma \Rightarrow P(|X-\mu| \geq 2\sigma) \leq \frac{\sigma^2}{4\sigma^2} = \frac{1}{4}$
- Αν $t = 3\sigma \Rightarrow P(|X-\mu| \geq 3\sigma) \leq \frac{\sigma^2}{9\sigma^2} = \frac{1}{9}$

I.e., the further from the expected value, the less the probability

- **Probability Generator Functions**

Definition: $\Pi(z) = E[z^X] = \sum_{x=0}^{\infty} p(x)\cdot z^x$

- $EX = \Pi'(1)$
- $Var X = \Pi''(1) + \Pi'(1) - (\Pi'(1))^2$

- **Moment Generator Functions**

$M(t) = \sum_x f(x)e^{tx}$, if X discrete

· Moment Generator Functions

Definition: $M(t) = E(e^{tX}) \Rightarrow \begin{cases} M(t) = \sum_{x} f(x)e^{tx}, & \text{if } X \text{ discrete} \\ M(t) = \int_{-\infty}^{\infty} f(x)e^{tx}dx, & \text{if } X \text{ continuous} \end{cases}$

- $EX = M'(0)$
- $EX^2 = M''(0)$

Exercise 1

Tuesday, November 5, 2024 11:50 AM

Let a, b be constants and the pdf of the r.v. X is: $f(x) = \begin{cases} ax + bx^2, & 0 \le x \le 1 \\ 0, & otherwise \end{cases}$

If E(X)=0.6, calculate:
(α) the constants a and b
(β) the distribution function
(γ) P(X<0.5)
(δ) Var(X)

a) $$EX = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x(ax+bx^2)\,dx = a\int_0^1 x^2\,dx + b\int_0^1 x^3\,dx =$$

$$= a\left[\frac{x^3}{3}\right]_0^1 + b\left[\frac{x^4}{4}\right]_0^1 = \frac{a}{3} + \frac{b}{4} = 0.6 \Rightarrow \boxed{4a+3b=7.2}\ (1)$$

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \Rightarrow \int_0^1 (ax+bx^2)\,dx = 1 \Rightarrow a\int_0^1 x\,dx + b\int_0^1 x^2\,dx = 1 \Rightarrow$$

$$a\left[\frac{x^2}{2}\right]_0^1 + b\left[\frac{x^3}{3}\right]_0^1 = 1 \Rightarrow \frac{a}{2} + \frac{b}{3} = 1 \Rightarrow \boxed{3a+2b=6}\ (2)$$

$$\left.\begin{aligned} (1) &\xRightarrow{\times 3} 12a+9b=21.6 \\ (2) &\xRightarrow{\times(-4)} -12a-8b=-24 \end{aligned}\right\} \xRightarrow{(+)} \boxed{b=-2.4}\ (3)$$

$$(2) \xRightarrow{(3)} 3a + 2(-2.4) = 3a - 4.8 = 6 \Rightarrow 3a = 10.8 \Rightarrow \boxed{a=3.6}$$

So, $$\boxed{f(x) = \begin{cases} 3.6x - 2.4x^2, & 0 \le x \le 1 \\ 0, & otherwise \end{cases}}$$

x x x x

b) $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \int_{0}^{x} (3.6t - 2.4t^2)\,dt = 3.6\int_{0}^{x} t\,dt - 2.4\int_{0}^{x} t^2\,dt =$

$= 3.6\left[\frac{t^2}{2}\right]_0^x - 2.4\left[\frac{t^3}{3}\right]_0^x \Rightarrow F(x) = \frac{3.6}{2}\cdot x^2 - \frac{2.4}{3}\cdot x^3 \Rightarrow$

$\boxed{F(x) = 1.8x^2 - 0.8x^3}$

c) $P(X<0.5) = F(0.5) = 1.8\cdot 0.5^2 - 0.8\cdot 0.5^3 = 1.8\cdot 0.25 - 0.8\cdot 0.125 =$

$= 0.45 - 0.1 \Rightarrow \boxed{P(X<0.5) = 0.35}$

d) $\mathrm{Var}X = EX^2 - (EX)^2$ $\qquad E[g(X)] = \int_{-\infty}^{\infty} g(x)\cdot f(x)\,dx$

$EX^2 = \int_0^1 x^2(3.6x - 2.4x^2)\,dx = 3.6\int_0^1 x^3\,dx - 2.4\int_0^1 x^4\,dx =$

$= 3.6\left[\frac{x^4}{4}\right]_0^1 - 2.4\left[\frac{x^5}{5}\right]_0^1 = \frac{3.6}{4} - \frac{2.4}{5} \Rightarrow EX^2 = 0.9 - 0.48 \Rightarrow \boxed{EX^2 = 0.42}$

$\mathrm{Var}X = 0.42 - (0.6)^2 = 0.42 - 0.36 \Rightarrow \boxed{\mathrm{Var}X = 0.06}$

Exercise 2

Wednesday, November 6, 2024 12:10 PM

Assume a fair coin with 2 sides, K and Γ. We toss the coin until 5 we repeat the experiment 5 times or until we get a K for the first time.

a) What is the expected value and the variance of the number of repetitions?
b) What is the expected value and the variance of the number of Γ?

a)

Results	#tosses	# Γ	Probability
K	1	0	1/2
ΓK	2	1	1/4
ΓΓK	3	2	1/8
ΓΓΓK	4	3	1/16
ΓΓΓΓK	5	4	1/32
ΓΓΓΓΓ	5	5	1/32

X: r.v. counting the number of tosses.

$$EX = \sum_{x=1}^{5} x P(X=x) = 1\cdot\frac{1}{2} + 2\cdot\frac{1}{4} + 3\cdot\frac{1}{8} + 4\cdot\frac{1}{16} + 5\left(\frac{1}{32}+\frac{1}{32}\right) = \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{1}{4} + \frac{5}{16} =$$

$$= \frac{8+8+6+4+5}{16} \Rightarrow \boxed{EX = \frac{31}{16}}$$

$$VarX = EX^2 - (EX)^2$$

$$EX^2 = \sum_{x=1}^{5} x^2 P(X=x) = 1^2\cdot\frac{1}{2} + 2^2\cdot\frac{1}{4} + 3^2\cdot\frac{1}{8} + 4^2\cdot\frac{1}{16} + 5^2\left(\frac{1}{32}+\frac{1}{32}\right) =$$

$$= \frac{1}{2} + 1 + \frac{9}{8} + 1 + \frac{25}{16} \Rightarrow EX^2 = \frac{8+16+18+16+25}{16} \Rightarrow \boxed{EX^2 = \frac{83}{16}}$$

$$VarX = EX^2 - (EX)^2 = \frac{83}{16} - \left(\frac{31}{16}\right)^2 \Rightarrow \boxed{VarX = \frac{367}{256} \simeq 1.43}$$

b) Y: r.v. counting the number of Γ

b) Y: r.v. counting the number of Γ

$$EY = \sum_{y=0}^{5} y P(Y=y) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{8} + 3 \cdot \frac{1}{16} + 4 \cdot \frac{1}{32} + 5 \cdot \frac{1}{32} =$$

$$= \frac{1}{4} + \frac{1}{4} + \frac{3}{16} + \frac{1}{8} + \frac{5}{32} = \frac{8+8+6+4+5}{32} \Rightarrow \boxed{EY = \frac{31}{32}}$$

$$\mathrm{Var}\,Y = EY^2 - (EY)^2$$

$$EY^2 = \sum_{y=0}^{5} y^2 P(Y=y) = 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{8} + 3^2 \cdot \frac{1}{16} + 4^2 \cdot \frac{1}{32} + 5^2 \cdot \frac{1}{32} =$$

$$= \frac{1}{4} + \frac{1}{4} + \frac{9}{16} + \frac{1}{2} + \frac{25}{32} = \frac{8+8+18+16+25}{32} \Rightarrow \boxed{EY^2 = \frac{83}{32}}$$

$$\mathrm{Var}\,Y = \frac{83}{32} - \left(\frac{31}{32}\right)^2 \simeq \frac{1695}{1024} \Rightarrow \boxed{\mathrm{Var}\,Y \simeq 1.66}$$

Exercise 3

Tuesday, November 5, 2024 12:22 PM

Let the number of items that are produced by a company in a week is a r.v. with expected value equal to 50.

a) Calculate an upper bound on the probability that the number of items produced in a certain week is at least 75.
b) If the variance is 25, then calculate a lower bound on the probability the production of items in a week is between 40 and 60.

X: r.v. counting the number of items produced in a week

$EX = 50$

a) Markov Inequality $P(X \geq t) \leq \frac{EX}{t} \Rightarrow$

$$\Rightarrow P(X \geq 75) \leq \frac{50}{75} \Rightarrow \boxed{P(X \geq 75) \leq 0.6667}$$

b) Chebyshev Inequality: $P(|X-\mu| \geq t) \leq \frac{VarX}{t^2}$

$$P(40 < X < 60) = P(40-50 < X-50 < 60-50) = P(-10 < X-50 < 10) =$$

$$= P(-10 < X-\mu < 10) = P(|X-\mu| < 10) = 1 - P(|X-\mu| \geq 10) \quad (1)$$

$$P(|X-\mu| \geq 10) \leq \frac{25}{10^2} = \frac{25}{100} = \frac{1}{4} \Rightarrow -P(|X-\mu| \geq 10) \geq -\frac{1}{4} \Rightarrow$$

$$\Rightarrow 1 - P(|X-\mu| \geq 10) \geq 1 - \frac{1}{4} \xRightarrow{(1)} P(|X-\mu| < 10) \geq \frac{3}{4} \Rightarrow$$

$$\Rightarrow \boxed{P(40 < X < 60) \geq \frac{3}{4}}$$

Exercise 4

Wednesday, November 6, 2024 12:01 PM

If the r.v. X has $\mathrm{EX} = 3$ and $EX^2 = 13$, calculate:

a) A lower bound for $P(-2 < X < 8)$
b) The expected value and variance of the r.v. $Y = 3X + 5$
c) The expected value of $Z = (X-1)^2$

a) $VarX = EX^2 - (EX)^2 = 13 - 3^2 \Rightarrow \boxed{VarX = 4}$

Chebyshev Inequality: $P(|X-EX| \geq t) \leq \frac{VarX}{t^2}$

$P(-2<X<8) = P(-5 < X-EX < 5) = P(|X-EX|<5) = 1 - P(|X-EX| \geq 5)$

$P(|X-EX| \geq 5) \leq \frac{4}{25} \Rightarrow 1 - P(|X-EX| \geq 5) \geq 1 - \frac{4}{25} \Rightarrow \boxed{P(|X-EX|<5) \geq \frac{21}{25}}$

b) $EY = E[3X+5] = 3EX + 5 = 3 \cdot 3 + 5 \Rightarrow \boxed{EY = 14}$

$VarY = Var[3X+5] = 3^2 \cdot VarX = 9 \cdot 4 \Rightarrow \boxed{VarY = 36}$

$\sigma = \sqrt{VarY} \Rightarrow \boxed{\sigma = 6}$

c) $EZ = E[X-1]^2 = E[X^2 - 2X + 1] = EX^2 - 2EX + 1 =$

$= 13 - 2 \cdot 3 + 1 \Rightarrow \boxed{EZ = 8}$

Exercise 5

Tuesday, November 5, 2024 1:10 PM

A non-negative integer random variable's X Moment Generator Functions is one of the following:

1) $M(t) = e^{2\left(e^{(e^t-1)}-1\right)}$

2) $M(t) = e^{2\left(e^{e^t}-1\right)}$

α) Which is the correct one?

β) Calculate P(X=0).

a) $M(t) = \sum_{x=0}^{\infty} P(x) e^{tx} \xRightarrow{t=0} M(0) = \sum_{x=0}^{\infty} P(x) e^{0\cdot x} \Rightarrow$

$\Rightarrow M(0) = \sum_{x=0}^{\infty} p(x) \Rightarrow \boxed{M(0) = 1}$

1) $t=0$: $M(0) = e^{2(e^{(e^0-1)}-1)} = e^{2(e^{(1-1)}-1)} = e^{2(1-1)} = e^0 \Rightarrow \boxed{M(0) = 1}$

2) $t=0$: $M(0) = e^{2(e^{e^0}-1)} = e^{2(e^1-1)} = e^{2(e-1)} \neq 1$

Thus, the correct is (1)

b) Let $e^t = z \Rightarrow t = \ln z$

$M(t) = M(\ln z) = \sum_{x=0}^{\infty} P(X=x)\cdot e^{(\ln z\cdot x)} = \sum_{x=0}^{\infty} P(X=x)\cdot\left(e^{\ln z}\right)^x =$

$= \sum_{x=0}^{\infty} P(X=x)\, z^x = \Pi(z)$

$\Pi(z) = P(X=0)\cdot z^0 + \sum_{x=1}^{\infty} P(X=x)\cdot z^x = P(X=0) + \sum_{x=1}^{\infty} P(X=x)\cdot z^x \Rightarrow$

$\xRightarrow{z=0} \boxed{\Pi(0) = P(X=0)}$

$$z = e^t \Rightarrow \lim_{t \to -\infty} e^t = 0$$

$$P(X=0) = \lim_{t \to -\infty} M(t) = \lim_{t \to -\infty} e^{2(e^{(e^t-1)}-1)} = e^{2(e^{-1}-1)}$$