Lecture 21:

CoVariance, Total Expectation

Lecture 20 Summary

- $Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$ measures the expected deviation squared from the mean.
- $\sigma(X) = \sqrt{Var(X)}$
- $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
- Two r.v.s X, Y are independent if
  for all a, b $\Pr[X=a, Y=b] = \Pr[X=a] \cdot \Pr[Y=b]$.
- If X, Y are independent then $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$
- If X, Y are independent, then $Var(X+Y) = Var(X) + Var(Y)$
- If $X_1, \ldots, X_n$ are pairwise indep. then
  $Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n)$
- If $X \sim Bin(n,p)$ then $Var(X) = n \cdot p \cdot (1-p)$.

Covariance

If X and Y are r.v.s their co-variance is

$$\mathrm{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}X)(Y-\mathbb{E}Y)\right]$$

measures whether greater values of X mainly correspond with greater values of Y.

Covariance

If X and Y are r.v.s their co-variance is

$$\mathrm{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}X)(Y-\mathbb{E}Y)\right]$$

measures whether greater values of X mainly correspond with greater values of Y.

Facts:

$$\mathrm{Cov}(X+a, Y+b) = \mathrm{Cov}(X,Y)$$

$$\mathrm{Cov}(c \cdot X, d \cdot Y) = c \cdot d \cdot \mathrm{Cov}(X,Y).$$

Covariance

If X and Y are r.v.s their co-variance is

$$\mathrm{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}X)(Y-\mathbb{E}Y)\right]$$

Fact: $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - (\mathbb{E}X)\cdot(\mathbb{E}Y)$

Covariance

If X and Y are r.v.s their co-variance is

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)]$$

Fact: $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - (\mathbb{E}X)\cdot(\mathbb{E}Y)$

Proof: let $\mu_x = \mathbb{E}X \qquad \mu_y = \mathbb{E}Y$

$$\begin{aligned}\mathrm{cov}(X,Y) &= \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] \\ &= \mathbb{E}[XY] - \mu_X\mathbb{E}[Y] - \mu_Y\mathbb{E}[X] + \mu_x\mu_y \\ &= \mathbb{E}[XY] - \mu_x\mu_y = \mathbb{E}[XY] - (\mathbb{E}X)(\mathbb{E}Y).\end{aligned}$$

Covariance

If X and Y are r.v.s their co-variance is

$$\mathrm{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}X)(Y-\mathbb{E}Y)\right]$$

Fact: $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - (\mathbb{E}X)\cdot(\mathbb{E}Y)$

Corollary: If X and Y are independent then

$$\mathrm{Cov}(X,Y) = 0.$$

Recap

Theorem: If X and Y are independent r.v.s, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Recap

Theorem: If X and Y are independent r.v.s, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Proof:

$$\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2]$$

$$(\mathbb{E}[X+Y])^2 = (\mathbb{E}[X] + \mathbb{E}[Y])^2$$

$$= (\mathbb{E}X)^2 + 2(\mathbb{E}X)(\mathbb{E}Y) + (\mathbb{E}Y)^2$$

Recap

Theorem: If X and Y are independent r.v.s, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Proof:

$$-\begin{cases} \mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2] \\ (\mathbb{E}[X+Y])^2 = (\mathbb{E}[X] + \mathbb{E}[Y])^2 \\ \qquad = (\mathbb{E}X)^2 + 2(\mathbb{E}X)(\mathbb{E}Y) + (\mathbb{E}Y)^2 \end{cases}$$

$$\mathrm{Var}(X+Y) = \mathbb{E}[X^2] - (\mathbb{E}X)^2 + 2\mathbb{E}[XY] - 2(\mathbb{E}X)(\mathbb{E}Y) + \mathbb{E}[Y^2] - (\mathbb{E}Y)^2$$

$$= \mathrm{Var}(X) + 0 + \mathrm{Var}(Y)$$

The only point where we used that X & Y are indep

Theorem: If X and Y are ~~independent~~ r.v.s, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\cdot\mathrm{Cov}(X,Y)$$

Proof:

$$-\begin{cases} \mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2] + 2\,\mathbb{E}[XY] + \mathbb{E}[Y^2] \\ (\mathbb{E}[X+Y])^2 = (\mathbb{E}X + \mathbb{E}Y)^2 \\ \qquad = (\mathbb{E}X)^2 + 2(\mathbb{E}X)(\mathbb{E}Y) + (\mathbb{E}Y)^2 \end{cases}$$

$$\begin{aligned} \mathrm{Var}(X+Y) &= \mathbb{E}[X^2] - (\mathbb{E}X)^2 + 2\,\mathbb{E}[XY] - 2(\mathbb{E}X)(\mathbb{E}Y) \\ &\quad + \mathbb{E}[Y^2] - (\mathbb{E}Y)^2 \\ &= \mathrm{Var}(X) + 2\,\mathrm{Cov}(X,Y) + \mathrm{Var}(Y) \end{aligned}$$

Variance of a Poisson RV

$X \sim \text{Pois}(\lambda) \quad \Pr[X=i] = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$

We saw: $\mathbb{E}X = \lambda$

$\text{Var}(X) = ?$

Variance of a Poisson RV

$X \sim \mathrm{Pois}(\lambda) \qquad \Pr[X=i] = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$

We saw: $\mathbb{E}X = \lambda$

$\mathrm{Var}(X) = ?$

Recall: $\mathrm{Bin}\left(n, \frac{\lambda}{n}\right) \xrightarrow[n \to \infty]{} \mathrm{Pois}(\lambda)$

Variance of a Poisson RV

$X \sim Pois(\lambda) \quad Pr[X=i] = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$

We saw: $\mathbb{E}X = \lambda$

$Var(X) = ?$

Recall:

$$Bin(n, \tfrac{\lambda}{n}) \xrightarrow[n\to\infty]{} Pois(\lambda)$$

$Y \sim Bin(n, \frac{\lambda}{n})$

$\mathbb{E}[Y] = n \cdot \frac{\lambda}{n} = \lambda = \mathbb{E}[X]$.

$Var(Y) = n \cdot \frac{\lambda}{n} \cdot (1 - \frac{\lambda}{n}) \xrightarrow[n\to\infty]{} \lambda = Var[X]$.

$\mathbb{E}X^2 = ?$

Normalizing a Random Variable

If X is a r.v. with mean μ and stdev σ

Then $\tilde{X} = \frac{X-\mu}{\sigma}$ is the normalized version of X.

$\mathbb{E}[\tilde{X}] = ?$

$\mathrm{Var}(\tilde{X}) = ?$

Normalizing a Random Variable

If X is a r.v. with mean μ and stdev σ

Then $\tilde{X} = \frac{X-\mu}{\sigma}$ is the normalized version of X.

$$\mathbb{E}[\tilde{X}] = \mathbb{E}\left[\frac{X-\mu}{\sigma}\right] = \frac{\mathbb{E}[X-\mu]}{\sigma} = 0$$

$$\mathrm{Var}(\tilde{X}) = \mathbb{E}[\tilde{X}^2] = \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^2\right] = \frac{\mathbb{E}[(X-\mu)^2]}{\sigma^2}$$

$$= \frac{\mathrm{Var}(X)}{\mathrm{Var}(X)} = 1.$$

Correlation

Def'n: The correlation of X, Y is defined as

$$\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma(X)\,\sigma(Y)}$$

Correlation

Def'n: The correlation of X, Y is defined as

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\,\sigma(Y)}$$

Claim: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X} \cdot \tilde{Y}]$

Correlation

Def'n: The correlation of X, Y is defined as

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

Claim: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

Proof:

$$\tilde{X} = \frac{X-\mu(X)}{\sigma(X)} \qquad \tilde{Y} = \frac{Y-\mu(Y)}{\sigma(Y)}$$

$$\mathbb{E}[\tilde{X}\cdot\tilde{Y}] = \mathbb{E}\left[\left(\frac{X-\mu(X)}{\sigma(X)}\right)\cdot\left(\frac{Y-\mu(Y)}{\sigma(Y)}\right)\right] = \frac{\mathbb{E}[(X-\mu(X))(Y-\mu(Y))]}{\sigma(X)\sigma(Y)}$$

$$= \mathrm{corr}(X,Y).$$

Def'n: The **Correlation** of X, Y is defined as

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\,\sigma(Y)}$$

Claim: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X} \cdot \tilde{Y}]$

Theorem: $-1 \leq \mathrm{Corr}(X,Y) \leq 1$ and equality holds iff $Y = \alpha X + \beta$.

Def'n: The <u>Correlation</u> of X, Y is defined as

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

Claim: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

Theorem: $-1 \le \mathrm{Corr}(X,Y) \le 1$ and equality holds iff $Y = \alpha X + \beta$.

Proof: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

(i) $0 \le \mathbb{E}[(\tilde{X}-\tilde{Y})^2] = \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] - 2\,\mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

$\therefore\ 2\,\mathbb{E}[\tilde{X}\tilde{Y}] \le \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] = 2.$ Hence, $\mathbb{E}[\tilde{X}\cdot\tilde{Y}] \le 1.$

Def'n: The <u>Correlation</u> of X, Y is defined as

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

Claim: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X} \cdot \tilde{Y}]$

Theorem: $-1 \leq \mathrm{Corr}(X,Y) \leq 1$ and equality holds iff $Y = \alpha X + \beta$.

Proof: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X} \cdot \tilde{Y}]$

(i) $0 \leq \mathbb{E}[(\tilde{X} - \tilde{Y})^2] = \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] - 2\mathbb{E}[\tilde{X} \cdot \tilde{Y}]$

$\therefore 2\mathbb{E}[\tilde{X}\tilde{Y}] \leq \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] = 2.$ Hence, $\mathbb{E}[\tilde{X} \cdot \tilde{Y}] \leq 1.$

(ii) $0 \leq \mathbb{E}[(\tilde{X} + \tilde{Y})^2] = \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] + 2\mathbb{E}[\tilde{X} \cdot \tilde{Y}]$

$\therefore 2\mathbb{E}[\tilde{X} \cdot \tilde{Y}] \geq -2$. Hence $\mathbb{E}[\tilde{X} \cdot \tilde{Y}] \geq -1.$

Def'n: The <u>Correlation</u> of X, Y is defined as

$$\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

Claim: $\text{Corr}(X,Y) = \mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

Theorem: $-1 \le \text{Corr}(X,Y) \le 1$ and equality holds iff $Y = \alpha X + \beta$.

Proof: $\text{Corr}(X,Y) = \mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

(i) $0 \le \mathbb{E}[(\tilde{X}-\tilde{Y})^2] = \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] - 2\,\mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

$\therefore\ 2\,\mathbb{E}[\tilde{X}\tilde{Y}] \le \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] = 2$. Hence, $\mathbb{E}[\tilde{X}\cdot\tilde{Y}] \le 1$.

(ii) $0 \le \mathbb{E}[(\tilde{X}+\tilde{Y})^2] = \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] + 2\,\mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

$\therefore\ 2\,\mathbb{E}[\tilde{X}\cdot\tilde{Y}] \ge -2$. Hence $\mathbb{E}[\tilde{X}\cdot\tilde{Y}] \ge -1$.

$$\text{Corr}(X,Y) = 1 \iff \mathbb{E}[(\tilde{X}-\tilde{Y})^2] = 0 \iff \tilde{X} = \tilde{Y} \iff \frac{X-\mu_X}{\sigma_X} = \frac{Y-\mu_Y}{\sigma_Y}$$

Def'n: The <u>Correlation</u> of X, Y is defined as

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

Claim: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

Theorem: $-1 \le \mathrm{Corr}(X,Y) \le 1$ and equality holds iff $Y = \alpha X + \beta$.

Proof: $\mathrm{Corr}(X,Y) = \mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

(i) $0 \le \mathbb{E}[(\tilde{X}-\tilde{Y})^2] = \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] - 2\mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

$\therefore\ 2\mathbb{E}[\tilde{X}\tilde{Y}] \le \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] = 2.$ Hence, $\mathbb{E}[\tilde{X}\cdot\tilde{Y}] \le 1.$

(ii) $0 \le \mathbb{E}[(\tilde{X}+\tilde{Y})^2] = \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] + 2\mathbb{E}[\tilde{X}\cdot\tilde{Y}]$

$\therefore\ 2\mathbb{E}[\tilde{X}\cdot\tilde{Y}] \ge -2$. Hence $\mathbb{E}[\tilde{X}\cdot\tilde{Y}] \ge -1.$

$\mathrm{Corr}(X,Y) = 1 \iff \mathbb{E}[(\tilde{X}-\tilde{Y})^2] = 0 \iff \tilde{X} = \tilde{Y} \iff \dfrac{X-\mu_X}{\sigma_X} = \dfrac{Y-\mu_Y}{\sigma_Y}$

$\mathrm{Corr}(X,Y) = -1 \iff \mathbb{E}[(\tilde{X}+\tilde{Y})^2] = 0 \iff \tilde{X} = -\tilde{Y} \iff \dfrac{X-\mu_X}{\sigma_X} = -\dfrac{Y-\mu_Y}{\sigma_Y}$

Example

X, Y are two r.v.s whose joint dist is

$Y \backslash X$	-1	0	$+1$
$+1$		$\frac{1}{5}$	
0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
-1		$\frac{1}{5}$	

Example

X, Y are two r.v.s whose joint dist is

$Y \setminus X$	-1	0	$+1$
$+1$		$\frac{1}{5}$	
0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
-1		$\frac{1}{5}$	

Q: Are X and Y independent?

Example

X, Y are two r.v.s whose joint dist is

$Y \backslash X$	-1	0	$+1$
$+1$		$\frac{1}{5}$	
0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
-1		$\frac{1}{5}$	

Q: Are X and Y independent?

Q: What's the $\mathrm{Cov}(X, Y)$?

Example

X, Y are two r.v.s whose joint dist is

Y \ X	-1	0	$+1$
$+1$		$\frac{1}{5}$	
0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
-1		$\frac{1}{5}$	

Q: Are X and Y independent?

Q: What's the $\mathrm{Cov}(X,Y)$?

$\mathbb{E}[XY] =$ $\qquad$ $\mathbb{E}[X] =$ $\qquad$ $\mathbb{E}[Y] =$

Independent $\Longrightarrow$ Uncorellated

but

Independent $\not\Longleftarrow$ Uncorellated

Example

X, Y are two r.v.s whose joint dist is

$Y \backslash X$	-1	0	$+1$
$+1$		$1/7$	$1/7$
0	$1/7$	$1/7$	$1/7$
-1	$1/7$	$1/7$	

Q: Are X and Y independent?

Q: What's the $\mathrm{Cov}(X, Y)$?

$\mathbb{E}[XY] =$ $\mathbb{E}[X] =$ $\mathbb{E}[Y] =$

Conditioning on Events

Let X be a r.v. over Ω

and let A be an event.

$$\Pr[X = x \mid A] =$$

Conditioning on Events

Let X be a r.v. over Ω

and let A be an event.

$$\Pr[X=x \mid A] = \frac{\Pr[(X=x)\cap A]}{\Pr[A]}$$

If $A_1, \ldots, A_n$ are disjoint events that partitions Ω

Then,

$\Pr[X=x] =$

Conditioning on Events

Let X be a r.v. over Ω and let A be an event.

$$\Pr[X=x \mid A] = \frac{\Pr[(X=x)\cap A]}{\Pr[A]}$$

If $A_1, \ldots, A_n$ are disjoint events that partitions Ω

Then,

$$\Pr[X=x] = \Pr[(X=x)\cap A_1] + \cdots + \Pr[(X=x)\cap A_n]$$

$$= \Pr[A_1]\cdot\Pr[X=x \mid A_1] + \cdots + \Pr[A_n]\cdot\Pr[X=x \mid A_n].$$

Law of Total Expectation

Theorem: Let X be a r.v. over Ω

and $A_1, \ldots, A_n$ be disjoint events that partitions Ω

then $\mathbb{E}[X] = \sum_{i=1}^{n} \Pr[A_i] \cdot \mathbb{E}[X \mid A_i]$

Law of Total Expectation

Theorem: Let X be a r.v. over Ω

and $A_1, \dots, A_n$ be disjoint events that partitions Ω

then $\mathbb{E}[X] = \sum_{i=1}^{n} \Pr[A_i] \cdot \mathbb{E}[X \mid A_i]$

Proof: $\mathbb{E}[X] = \sum_{x} x \cdot \Pr[X = x]$

$= \sum_{x} x \cdot \sum_{i=1}^{n} \Pr[A_i] \cdot \Pr[X = x \mid A_i]$

$= \sum_{i=1}^{n} \Pr[A_i] \cdot \sum_{x} x \cdot \Pr[X = x \mid A_i]$

$= \sum_{i=1}^{n} \Pr[A_i] \cdot \mathbb{E}[X \mid A_i]$

Example: Computing the Expectation of a Geometric R.V.

$X \sim Geo(p)$

We saw using algebraic tricks that $\mathbb{E}X = \frac{1}{p}$

Random Experiment: toss a coin with heads prob. p until you get H.

Let's see another proof.

$X \sim Geo(p)$ $\Omega = \{H, TH, TTH, \ldots\}$

$Pr[X=i] = (1-p)^{i-1} \cdot p$ for $i = 1, 2, \ldots$

$X \sim Geo(p)$ $\qquad \Omega = \{H, TH, TTH, \ldots\}$

$Pr[X=i] = (1-p)^{i-1} \cdot p \qquad$ for $i = 1, 2, \ldots$

Idea: Partition the sample space to two parts:

A = first toss is heads, $\bar{A}$ = first toss is tails.

$X \sim Geo(p)$ $\Omega = \{H, TH, TTH, \ldots\}$

$Pr[X=i] = (1-p)^{i-1} \cdot p$ for $i=1,2,\ldots$

Idea: Partition the sample space to two parts:

A = first toss is heads, $\bar{A}$ = first toss is tails.

$Pr[A] = p$ $Pr[\bar{A}] = 1-p$.

$X|A = 1$ $X|\bar{A} = \left\{\begin{array}{ll} 2, & \text{w.p. } p \\ 3, & \text{w.p. } p\cdot(1-p) \\ \vdots & \end{array}\right\} = 1+Y$ where $Y \sim Geo(p)$

$X \sim Geo(p) \qquad \Omega = \{H, TH, TTH, \dots\}$

$\Pr[X=i] = (1-p)^{i-1} \cdot p \qquad$ for $i = 1, 2, \dots$

Idea: Partition the sample space to two parts:

A = first toss is heads, $\bar{A}$ = first toss is tails.

$\Pr[A] = p \qquad \Pr[\bar{A}] = 1-p$

$$X|A = 1 \qquad X|\bar{A} = \left\{\begin{array}{ll} 2, & \text{w.p. } p \\ 3, & \text{w.p. } p\cdot(1-p) \\ \vdots & \end{array}\right\} = 1+Y$$

where $Y \sim Geo(p)$

$$\mathbb{E}[X] = \Pr[A]\cdot\mathbb{E}[X|A] + \Pr[\bar{A}]\cdot\mathbb{E}[X|\bar{A}]$$

$$= p\cdot 1 + (1-p)\cdot\mathbb{E}[1+Y] = p\cdot 1 + (1-p)\cdot\left(1+\mathbb{E}[Y]\right) = 1 + (1-p)\cdot\mathbb{E}[X]$$

$X \sim Geo(p) \qquad \Omega = \{H, TH, TTH, \dots\}$

$Pr[X=i] = (1-p)^{i-1} \cdot p \qquad$ for $i=1,2,\dots$

Idea: Partition the sample space to two parts:

A = first toss is heads, $\bar{A}$ = first toss is tails.

$Pr[A] = p \qquad Pr[\bar{A}] = 1-p$.

$$X|A = 1 \qquad X|\bar{A} = \left\{\begin{array}{ll} 2, & \text{w.p. } p \\ 3, & \text{w.p. } p\cdot(1-p) \\ \vdots & \end{array}\right\} = 1+Y$$

where $Y \sim Geo(p)$

$$\mathbb{E}[X] = Pr[A]\cdot\mathbb{E}[X|A] + Pr[\bar{A}]\cdot\mathbb{E}[X|\bar{A}]$$

$$= p\cdot 1 + (1-p)\cdot\mathbb{E}[1+Y] = p\cdot 1 + (1-p)\cdot(1+\mathbb{E}[Y]) = 1 + (1-p)\cdot\mathbb{E}[X]$$

Hence $p\cdot\mathbb{E}X = 1$ and we get $\mathbb{E}[X] = \frac{1}{p}$.

Computing the Variance of a Geometric RV

$X \sim Geo(p)$ $\qquad \Omega = \{H, TH, TTH, \dots\}$

$\Pr[X=i] = (1-p)^{i-1} \cdot p \qquad$ for $i=1,2,\dots$

Idea: Partition the sample space to two parts:

A = first toss is heads, $\bar{A}$ = first toss is tails.

$\Pr[A] = p \qquad \Pr[\bar{A}] = 1-p.$

$X|A = 1 \qquad X|\bar{A} = 1+Y \qquad$ where $\qquad Y \sim Geo(p)$

$\mathbb{E}[X^2] = ?$

Computing the Variance of a Geometric RV

$X \sim Geo(p)$ $\quad \Omega = \{H, TH, TTH, \ldots\}$

$Pr[X=i] = (1-p)^{i-1} \cdot p$ $\quad$ for $i = 1, 2, \ldots$

Idea: Partition the sample space to two parts:

A = first toss is heads, $\bar{A}$ = first toss is tails.

$Pr[A] = p$ $\quad Pr[\bar{A}] = 1-p.$

$X|A = 1$ $\quad X|\bar{A} = 1+Y$ $\quad$ where $\quad Y \sim Geo(p)$

$$\mathbb{E}[X^2] = Pr[A] \cdot \mathbb{E}[X^2|A] + Pr[\bar{A}] \cdot \mathbb{E}[X^2|\bar{A}]$$

$$= p \cdot 1 + (1-p) \cdot \mathbb{E}[(1+Y)^2]$$

$$= p + (1-p) \cdot (1 + 2\mathbb{E}[Y] + \mathbb{E}[Y^2])$$

$$= 1 + \frac{(1-p) \cdot 2}{p} + (1-p)\,\mathbb{E}[X^2]$$

Computing the Variance of a Geometric RV

$X \sim Geo(p)$ $\Omega = \{H, TH, TTH, \dots\}$

$Pr[X=i] = (1-p)^{i-1} \cdot p$ for $i = 1, 2, \dots$

Idea: Partition the sample space to two parts:

A = first toss is heads, $\bar{A}$ = first toss is tails.

$Pr[A] = p$ $Pr[\bar{A}] = 1-p$.

$X|A = 1$ $X|\bar{A} = 1+Y$ where $Y \sim Geo(p)$

$$\mathbb{E}[X^2] = Pr[A] \cdot \mathbb{E}[X^2|A] + Pr[\bar{A}] \cdot \mathbb{E}[X^2|\bar{A}]$$

$$= p \cdot 1 + (1-p) \cdot \mathbb{E}[(1+Y)^2]$$

$$= p + (1-p) \cdot (1 + 2\mathbb{E}[Y] + \mathbb{E}[Y^2])$$

$$= 1 + \frac{(1-p) \cdot 2}{p} + (1-p)\mathbb{E}[X^2]$$

$$\therefore \mathbb{E}[X^2] \cdot p = \frac{2-p}{p}$$

$$\therefore \mathbb{E}[X^2] = \frac{2-p}{p^2}$$

$$\therefore Var[X] = \frac{1-p}{p^2}$$

Law of Iterated Expectation

Total Expectation: Let X be a r.v. over Ω

and $A_1, \ldots, A_n$ be disjoint events that partitions Ω

then $\mathbb{E}[X] = \sum_{i=1}^{n} \Pr[A_i] \cdot \mathbb{E}[X \mid A_i]$

Special Case: Let X, Y be r.v.s over Ω, then

$$\mathbb{E}[X] = \sum_{y \in \text{range}(Y)} \Pr[Y=y] \cdot \mathbb{E}[X \mid Y=y]$$

Law of Iterated Expectation - Example

You pick $N \sim Pois(\lambda)$.

Then you a coin w. heads prob p, N times.

X - the number of heads you got.

Calculate: $\mathbb{E}[X]$

Law of Iterated Expectation - Example

You pick $N \sim Pois(\lambda)$.

Then you a coin w. heads prob p, N times.

X - the number of heads you got.

Calculate: $\mathbb{E}[X]$

Solution: $\mathbb{E}[X] = \sum_{n=1}^{\infty} \Pr[N=n] \cdot \mathbb{E}[X \mid N=n]$

$$= \sum_{n=1}^{\infty} \Pr[N=n] \cdot p \cdot n = \mathbb{E}[p \cdot N] = p \cdot \mathbb{E}[N] = p \cdot \lambda.$$