Alex Psomas: Lecture 17.

Random Variables: Variance

1. Variance
2. Distributions

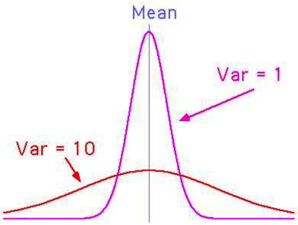
Variance

Flip a coin: If H you make a dollar. If T you lose a dollar.
Let X be the RV indicating how much money you make.
$E(X) = 0$.

Flip a coin: If H you make a million dollars. If T you lose a million dollars.
Let Y be the RV indicating how much money you make.
$E(Y) = 0$.

Any other measures??? What else that's informative can we say?

Variance



The variance measures the deviation from the mean value.

Definition: The variance of X is

$$\sigma^2(X) := var[X] = E[(X - E[X])^2].$$

$\sigma(X)$ is called the standard deviation of X.

Variance and Standard Deviation

Fact:

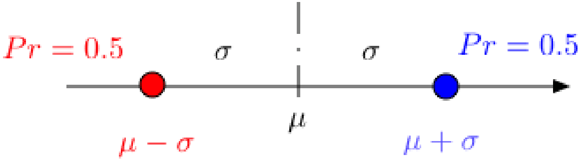
$$var[X] = E[X^2] - E[X]^2.$$

Indeed:

$$\begin{aligned}
var(X) &= E[(X - E[X])^2] \\
&= E[X^2 - 2XE[X] + E[X]^2 \\
&= E[X^2] - E[2XE[X]] + E[E[X]^2] \text{ by linearity} \\
&= E[X^2] - 2E[X]E[X] + E[X]^2, \\
&= E[X^2] - E[X]^2.
\end{aligned}$$

A simple example

This example illustrates the term ‘standard deviation.’



Consider the random variable X such that

$$X = \begin{cases} \mu - \sigma, & \text{w.p. } 1/2 \\ \mu + \sigma, & \text{w.p. } 1/2. \end{cases}$$

Then, $E[X] = \mu$ and $(X - E[X])^2 = \sigma^2$. Hence,

$$var(X) = \sigma^2 \text{ and } \sigma(X) = \sigma.$$

Example

Consider X with

$$X = \begin{cases} -1, & \text{w. p. } 0.99 \\ 99, & \text{w. p. } 0.01. \end{cases}$$

Then

$$\begin{aligned} E[X] &= -1 \times 0.99 + 99 \times 0.01 = 0. \\ E[X^2] &= 1 \times 0.99 + (99)^2 \times 0.01 \approx 100. \\ Var(X) &\approx 100 \implies \sigma(X) \approx 10. \end{aligned}$$

Properties of variance.

1. $Var(cX) = c^2 Var(X)$, where c is a constant.
   Scales by c^2.
2. $Var(X + c) = Var(X)$, where c is a constant.
   Shifts center.

Proof:

$$\begin{aligned}
Var(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2) \\
&= c^2 Var(X) \\
Var(X + c) &= E((X + c - E(X + c))^2) \\
&= E((X + c - E(X) - c)^2) \\
&= E((X - E(X))^2) = Var(X)
\end{aligned}$$

$\square$

Variance of sum of two independent random variables

Theorem:
If X and Y are independent, then

$$Var(X+Y) = Var(X) + Var(Y).$$

Proof:
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$

Hence,

$$\begin{aligned}
var(X+Y) &= E((X+Y)^2) = E(X^2 + 2XY + Y^2) \\
&= E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2) \\
&= E(X^2) - (E(X))^2 + E(Y^2) - (E(Y))^2 = var(X) + var(Y).
\end{aligned}$$

Variance of sum of independent random variables

Theorem:
If $X, Y, Z, \ldots$ are pairwise independent, then

$$var(X + Y + Z + \cdots) = var(X) + var(Y) + var(Z) + \cdots.$$

Proof:
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

$$E[XY] = E[X]E[Y] = 0. \text{ Also, } E[XZ] = E[YZ] = \cdots = 0.$$

Hence,

$$\begin{aligned}
var(X + Y + Z + \cdots) &= E((X + Y + Z + \cdots)^2) \\
&= E(X^2 + Y^2 + Z^2 + \cdots + 2XY + 2XZ + 2YZ + \cdots) \\
&= E(X^2) + E(Y^2) + E(Z^2) + \cdots + 0 + \cdots + 0 \\
&= var(X) + var(Y) + var(Z) + \cdots.
\end{aligned}$$

$\square$

Distributions

- Bernoulli
- Binomial
- Uniform
- Geometric
- Poisson

Bernoulli

Flip a coin, with heads probability p.

Random variable X: 1 is heads, 0 if not heads.

X has the Bernoulli distribution.

Distribution:

$$X = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1^2 \times p + 0^2 \times (1-p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1-p)$$

Binomial

Flip n coins with heads probability p.

Random variable: number of heads.

Binomial Distribution: $Pr[X = i]$, for each i.

How many sample points in event "$X = i$"?
i heads out of n coin flips $\implies \binom{n}{i}$

Sample space: $\Omega = \{HHH...HH, HHH...HT, \dots\}$
What is the probability of ω if ω has i heads?
Probability of heads in any position is p.
Probability of tails in any position is $(1-p)$.
So, we get $Pr[\omega] = p^i(1-p)^{n-i}$.
Probability of "$X = i$" is sum of $Pr[\omega]$, $\omega \in$ "$X = i$".

$$Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}, i = 0, 1, \dots, n : B(n,p) \text{ distribution}$$

Expectation of Binomial Distribution

$$X_i = \begin{cases} 1 & \text{if } i\text{th flip is heads} \\ 0 & \text{otherwise} \end{cases}$$

$E[X_i] = 1 \times Pr[``heads''] + 0 \times Pr[``tails''] = p.$

Moreover $X = X_1 + \cdots X_n$ and

$E[X] = E[X_1] + E[X_2] + \cdots E[X_n] = n \times E[X_i] = np.$

Variance of Binomial Distribution.

Flip coin with heads probability p.
X- how many heads?

$$X_i = \begin{cases} 1 & \text{if } i\text{th flip is heads} \\ 0 & \text{otherwise} \end{cases}$$

$E(X_i^2) = 1^2 \times p + 0^2 \times (1-p) = p$.
$Var(X_i) = p - (E(X))^2 = p - p^2 = p(1-p)$.

$p = 0 \implies Var(X_i) = 0$
$p = 1 \implies Var(X_i) = 0$

$X = X_1 + X_2 + \ldots X_n$.

X_i and X_j are independent: $Pr[X_i = 1 | X_j = 1] = Pr[X_i = 1]$.

$$Var(X) = Var(X_1 + \cdots X_n) = np(1-p).$$

Uniform Distribution

Roll a six-sided balanced die. Let X be the number of pips (dots). Then X is equally likely to take any of the values $\{1, 2, \ldots, 6\}$. We say that X is *uniformly distributed* in $\{1, 2, \ldots, 6\}$.

More generally, we say that X is uniformly distributed in $\{1, 2, \ldots, n\}$ if $Pr[X = m] = 1/n$ for $m = 1, 2, \ldots, n$.
In that case,

$$E[X] = \sum_{m=1}^{n} m Pr[X = m] = \sum_{m=1}^{n} m \times \frac{1}{n} = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Variance of Uniform

Assume that $Pr[X = i] = 1/n$ for $i \in \{1, \ldots, n\}$. Then

$$\begin{aligned} E[X] &= \sum_{i=1}^{n} i \times Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i \\ &= \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}. \end{aligned}$$

Also,

$$\begin{aligned} E[X^2] &= \sum_{i=1}^{n} i^2 Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i^2 \\ &= \frac{1 + 3n + 2n^2}{6}, \text{ as you can verify.} \end{aligned}$$

This gives

$$var(X) = \frac{1 + 3n + 2n^2}{6} - \frac{(n+1)^2}{4} = \frac{n^2 - 1}{12}.$$

Geometric Distribution

Let's flip a coin with $Pr[H] = p$ until we get H.

For instance:

$$\begin{aligned}\omega_1 &= H, \text{ or}\\ \omega_2 &= T\ H, \text{ or}\\ \omega_3 &= T\ T\ H, \text{ or}\\ \omega_n &= T\ T\ T\ T\ \cdots\ T\ H.\end{aligned}$$

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$.

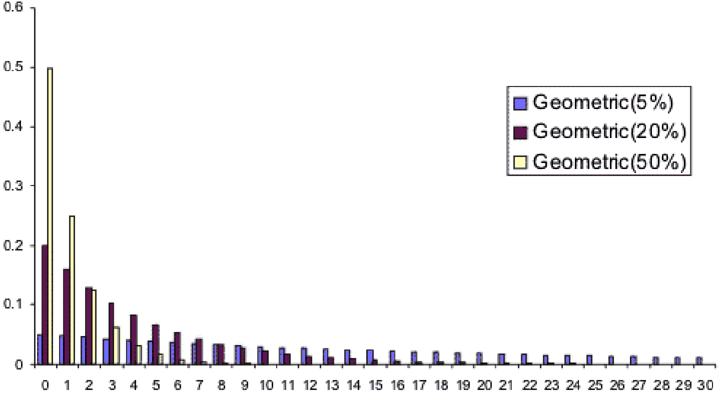
Let X be the number of flips until the first H. Then, $X(\omega_n) = n$.

Also,

$$Pr[X = n] = (1-p)^{n-1}p, \ n \geq 1.$$

Geometric Distribution

$$Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$



Geometric Distribution

$$Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

Note that

$$\sum_{n=1}^{\infty} Pr[X_n] = \sum_{n=1}^{\infty} (1-p)^{n-1}p = p\sum_{n=1}^{\infty} (1-p)^{n-1} = p\sum_{n=0}^{\infty} (1-p)^{n}.$$

Now, if $|a| < 1$, then $S := \sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$. Indeed,

$$\begin{aligned} S &= 1 + a + a^2 + a^3 + \cdots \\ aS &= \phantom{1 +{}} a + a^2 + a^3 + a^4 + \cdots \\ (1-a)S &= 1 + a - a + a^2 - a^2 + \cdots = 1. \end{aligned}$$

Hence,

$$\sum_{n=1}^{\infty} Pr[X_n] = p\,\frac{1}{1-(1-p)} = 1.$$

Geometric Distribution: Expectation

$$X =_D G(p), \text{ i.e., } Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

One has

$$E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1-p)^{n-1}p.$$

Thus,

$$\begin{aligned}
E[X] &= p + 2(1-p)p + 3(1-p)^2p + 4(1-p)^3p + \cdots \\
(1-p)E[X] &= \qquad (1-p)p + 2(1-p)^2p + 3(1-p)^3p + \cdots \\
pE[X] &= p + \ (1-p)p + \ (1-p)^2p + \ (1-p)^3p + \cdots \\
&\qquad \text{by subtracting the previous two identities} \\
&= \sum_{n=1}^{\infty} Pr[X = n] = 1.
\end{aligned}$$

Hence,

$$E[X] = \frac{1}{p}.$$

Coupon Collectors Problem.

Experiment: Get coupons at random from n until collect all n coupons.
Outcomes: $\{123145..., 56765...\}$
Random Variable: X - length of outcome.

Before: $Pr[X \geq n \ln 2n] \leq \frac{1}{2}$.

Today: $E[X]$?

Time to collect coupons

X-time to get n coupons.

X_1 - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

X_2 - time to get second coupon after getting first.

$Pr[\text{"get second coupon"}|\text{"got first coupon''}] = \frac{n-1}{n}$

$E[X_2]$? Geometric ! ! ! $\implies E[X_2] = \frac{1}{p} = \frac{1}{\frac{n-1}{n}} = \frac{n}{n-1}$.

$Pr[\text{"getting } i\text{th coupon}|\text{"got } i-1\text{rst coupons"}] = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

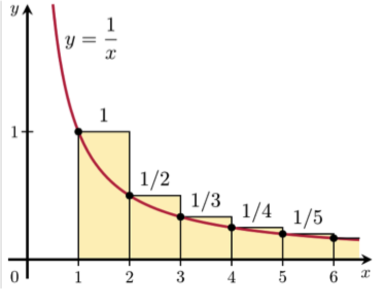
$E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}, i = 1, 2, \ldots, n.$

$$\begin{aligned} E[X] &= E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\ &= n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) =: nH(n) \approx n(\ln n + \gamma) \end{aligned}$$

Review: Harmonic sum

$$H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \int_1^n \frac{1}{x} dx = \ln(n).$$

.

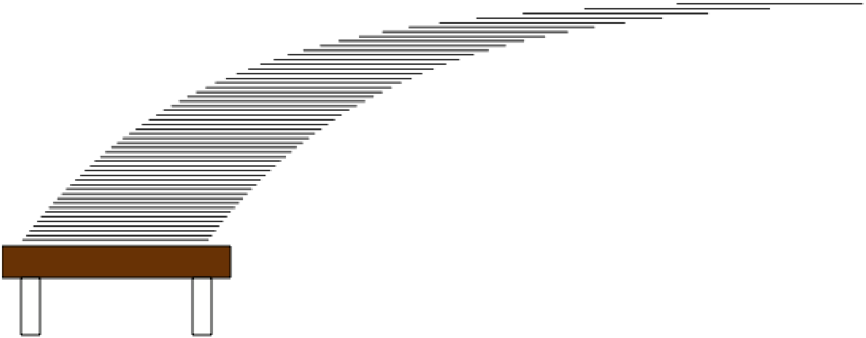


A good approximation is

$H(n) \approx \ln(n) + \gamma$ where $\gamma \approx 0.58$ (Euler-Mascheroni constant).

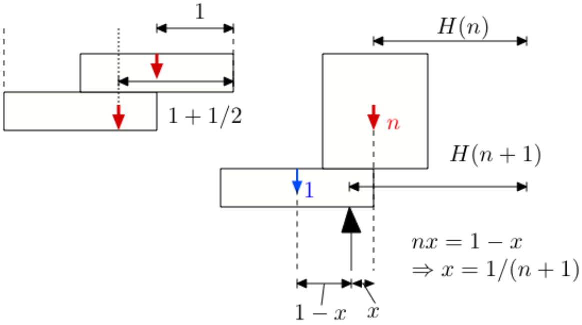
Harmonic sum: Paradox

Consider this stack of cards (no glue!):



If each card has length 2, the stack can extend $H(n)$ to the right of the table. As n increases, you can go as far as you want!

Stacking



The cards have width 2. Induction shows that the center of gravity after n cards is $H(n)$ away from the right-most edge.

Geometric Distribution: Memoryless

Let X be $G(p)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[\text{ first } n \text{ flips are } T] = (1-p)^n.$$

Theorem

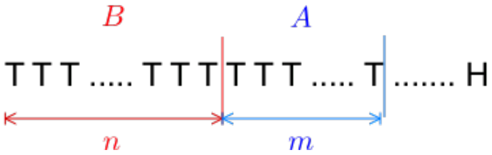
$$Pr[X > n+m | X > n] = Pr[X > m], m, n \geq 0.$$

Proof:

$$\begin{aligned}
Pr[X > n+m | X > n] &= \frac{Pr[X > n+m \text{ and } X > n]}{Pr[X > n]} \\
&= \frac{Pr[X > n+m]}{Pr[X > n]} \\
&= \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m \\
&= Pr[X > m].
\end{aligned}$$

Geometric Distribution: Memoryless - Interpretation

$$Pr[X > n+m | X > n] = Pr[X > m], m, n \geq 0.$$



$$Pr[X > n+m | X > n] = Pr[A|B] = Pr[A] = Pr[X > m].$$

The coin is memoryless, therefore, so is X.

Geometric Distribution: Yet another look

Theorem: For a r.v. X that takes the values $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

[See later for a proof.]

If $X = G(p)$, then $Pr[X \geq i] = Pr[X > i-1] = (1-p)^{i-1}$.

Hence,

$$E[X] = \sum_{i=1}^{\infty} (1-p)^{i-1} = \sum_{i=0}^{\infty} (1-p)^{i} = \frac{1}{1-(1-p)} = \frac{1}{p}.$$

Expected Value of Integer RV

Theorem: For a r.v. X that takes values in $\{0,1,2,\ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

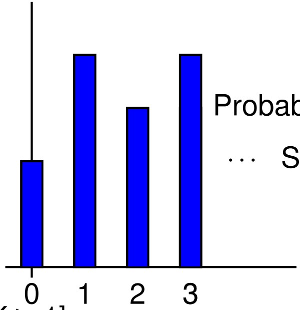
Proof: One has

$$\begin{aligned}
E[X] &= \sum_{i=1}^{\infty} i \times Pr[X = i] \\
&= \sum_{i=1}^{\infty} i\{Pr[X \geq i] - Pr[X \geq i+1]\} \\
&= \sum_{i=1}^{\infty} \{i \times Pr[X \geq i] - i \times Pr[X \geq i+1]\} \\
&= \sum_{i=1}^{\infty} \{i \times Pr[X \geq i] - (i-1) \times Pr[X \geq i]\} \\
&= \sum_{i=1}^{\infty} Pr[X \geq i].
\end{aligned}$$

□

Theorem: For a r.v. X that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$



Probability mass at i, counted i times.

$\cdots$ Same as $\sum_{i=1}^{\infty} i \times Pr[X = i]$.

$Pr[X \geq 1] \longrightarrow$

$Pr[X \geq 2] \longrightarrow$

$Pr[X \geq 3] \longrightarrow$

$\vdots$

Variance of geometric distribution.

X is a geometrically distributed RV with parameter p.
Thus, $Pr[X=n]=(1-p)^{n-1}p$ for $n \geq 1$. Recall $E[X]=1/p$.

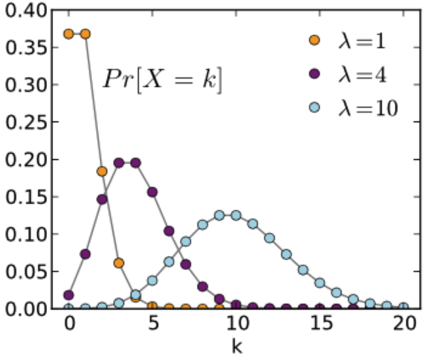
$$
\begin{aligned}
E[X^2] &= p+4p(1-p)+9p(1-p)^2+... \\
-(1-p)E[X^2] &= \quad -[p(1-p)+4p(1-p)^2+...] \\
pE[X^2] &= p+3p(1-p)+5p(1-p)^2+... \\
&= 2(p+2p(1-p)+3p(1-p)^2+..) \quad E[X]! \\
& \quad -(p+p(1-p)+p(1-p)^2+...) \quad \text{Distribution.} \\
pE[X^2] &= 2E[X]-1 \\
&= 2(\frac{1}{p})-1=\frac{2-p}{p}
\end{aligned}
$$

$\implies E[X^2]=(2-p)/p^2$ and
$var[X]=E[X^2]-E[X]^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}$.
$\sigma(X)=\frac{\sqrt{1-p}}{p} \approx E[X]$ when p is small(ish).

Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.

Poisson Distribution is distribution of X "for large n."



Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.
Poisson Distribution is distribution of X "for large n."
We expect $X \ll n$. For $m \ll n$ one has

$$
\begin{aligned}
Pr[X = m] &= \binom{n}{m} p^m (1-p)^{n-m}, \text{ with } p = \lambda/n \\
&= \frac{n(n-1)\cdots(n-m+1)}{m!} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&= \frac{n(n-1)\cdots(n-m+1)}{n^m} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&\approx^{(1)} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m} \approx^{(2)} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^n \approx \frac{\lambda^m}{m!} e^{-\lambda}.
\end{aligned}
$$

For (1) we used $m \ll n$; for (2) we used $(1 - a/n)^n \approx e^{-a}$.

Poisson Distribution: Definition and Mean

Definition Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \Leftrightarrow Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$

Fact: $E[X] = \lambda$.

Proof:

$$\begin{aligned} E[X] &= \sum_{m=1}^{\infty} m \times \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!} \\ &= e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^{m+1}}{m!} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \\ &= e^{-\lambda} \lambda e^{\lambda} = \lambda. \end{aligned}$$

$\square$

Review: Distributions

- $Bern(p) : Pr[X = 1] = p$;
  $E[X] = p$;
  $Var[X] = p(1-p)$;

- $B(n,p) : Pr[X = m] = \binom{n}{m} p^m (1-p)^{n-m}, m = 0, \ldots, n$;
  $E[X] = np$;
  $Var[X] = np(1-p)$;

- $U[1, \ldots, n] : Pr[X = m] = \frac{1}{n}, m = 1, \ldots, n$;
  $E[X] = \frac{n+1}{2}$;
  $Var[X] = \frac{n^2-1}{12}$;

- $G(p) : Pr[X = n] = (1-p)^{n-1} p, n = 1, 2, \ldots$;
  $E[X] = \frac{1}{p}$;
  $Var[X] = \frac{1-p}{p^2}$;

- $P(\lambda) : Pr[X = n] = \frac{\lambda^n}{n!} e^{-\lambda}, n \geq 0$;
  $E[X] = \lambda$;
  $Var[X] = \lambda$.

Today’s gig: Two envelopes problem.

Gigs so far:

1. How to tell random from human.
2. Monty Hall.
3. Birthday Paradox.
4. St. Petersburg paradox.
5. Simpson’s paradox.

Today:

Summary

$\boxed{\text{Random Variables}}$

- Variance.
- Distributions.