CS70: Lecture 26.

Recap of Distributions, Variance

1. Review: Distributions (Poisson)
2. Variance
3. Independence of Random Variables

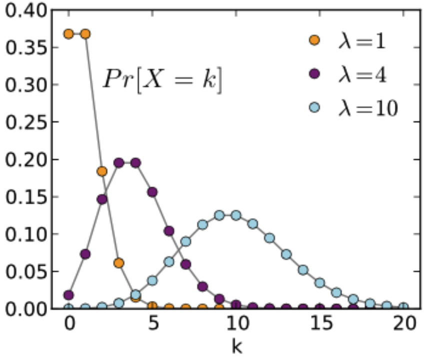
Review: Distributions

- $U[1,\ldots,n] : Pr[X = m] = \frac{1}{n}, m = 1,\ldots,n;$
  $E[X] = \frac{n+1}{2};$
- $B(n,p) : Pr[X = m] = \binom{n}{m} p^m (1-p)^{n-m}, m = 0,\ldots,n;$
  $E[X] = np;$
- $G(p) : Pr[X = n] = (1-p)^{n-1} p, n = 1,2,\ldots;$
  $E[X] = \frac{1}{p};$
- $P(\lambda) : Pr[X = n] = \frac{\lambda^n}{n!} e^{-\lambda}, n \geq 0;$
  $E[X] = \lambda.$

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."



The Poisson distribution shows up in a lot of “real world” applications. Here is a partial list:

- the number of bankruptcies that are filed in a month
- the number of arrivals at a car wash in one hour
- the number of arrivals at the Cory & Hearst bus-stop
- the number of network failures per day
- the number of asthma patient arrivals in a given hour at a walk-in clinic
- the number of customer arrivals at McDonald’s per day
- the number of birth, deaths, marriages, divorces, suicides, and homicides over a given period of time
- the number of visitors to a web site per minute

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

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

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

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:

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

$\square$

Simeon Poisson

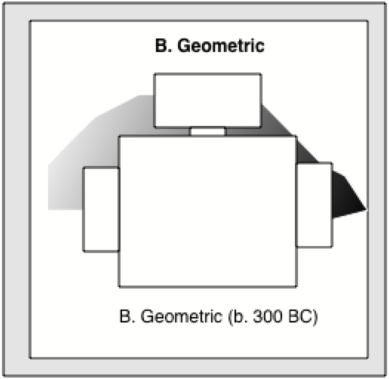
The Poisson distribution is named after:

Siméon Poisson

Siméon Denis Poisson (1781–1840)

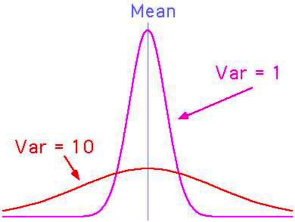
Equal Time: B. Geometric

The geometric distribution is named after:



I could not find a picture of D. Binomial, sorry.

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:

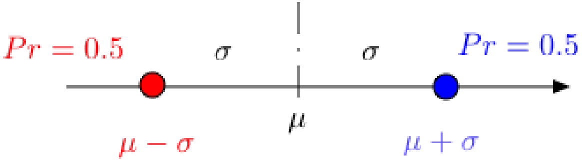
$$var[X] = E[X^2] - E[X]^2.$$

Indeed:

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

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

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

Also,

$$E(|X|) = 1 \times 0.99 + 99 \times 0.01 = 1.98.$$

Thus, $\sigma(X) \neq E[|X - E[X]|]$!

Exercise: How big can you make $\frac{\sigma(X)}{E[|X-E[X]|]}$?

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

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

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).

Fixed points.

Number of fixed points in a random permutation of n items.
"Number of student that get homework back."

$X = X_1 + X_2 \cdots + X_n$

where X_i is indicator variable for ith student getting hw back.

$$\begin{aligned} E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\ &= n \times \frac{1}{n} + (n)(n-1) \times \frac{1}{n(n-1)} \\ &= 1 + 1 = 2. \end{aligned}$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$
$= \frac{1}{n}$
$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{"anything else"}]$
$= \frac{1 \times 1 \times (n-2)!}{n!} = \frac{1}{n(n-1)}$

$Var(X) = E(X^2) - (E(X))^2 = 2 - 1 = 1.$

Variance: binomial.

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \\ &= \text{Really???!!\#\#...} \end{aligned}$$

Too hard!

Ok.. fine.
Let's do something else.
Maybe not much easier...but there is a payoff.

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:

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

$\square$

Independent random variables.

Definition: Independence

The random variables X and Y are **independent** if and only if

$$P[Y = b \mid X = a] = P[Y = b], \text{ for all } a \text{ and } b.$$

Fact:

X, Y are independent if and only if

$$P[X = a, Y = b] = P[X = a]P[Y = b], \text{ for all } a \text{ and } b.$$

Independence: examples.

Example 1
Roll two dices. X, Y = number of pips on the two dice. X, Y are independent.

Indeed: $P[X = a, Y = b] = 1/36, P[X = a] = P[Y = b] = 1/6$.

Example 2
Roll two dices. X = total number of pips, Y = number of pips on die 1 minus number on die 2. X and Y are not independent.

Indeed: $P[X = 12, Y = 1] = 0 \neq P[X = 12]P[Y = 1] > 0$.

Mean of product of independent RVs.

Theorem

Let X, Y be independent RVs. Then

$$E[XY] = E[X]E[Y].$$

Proof:

Recall that $E[g(X,Y)] = \sum_{x,y} g(x,y)P[X=x, Y=y]$. Hence,

$$\begin{aligned}
E[XY] &= \sum_{x,y} xyP[X=x, Y=y] = \sum_{x,y} xyP[X=x]P[Y=y], \text{ by ind.} \\
&= \sum_{x}\Big[\sum_{y} xyP[X=x]P[Y=y]\Big] \\
&= \sum_{x}\Big[xP[X=x]\Big(\sum_{y} yP[Y=y]\Big)\Big] \\
&= \sum_{x} xP[X=x]E[Y] = E[X]E[Y].
\end{aligned}$$

$\square$

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,

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

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: $Pr[X_i = 1 | X_j = 1] = Pr[X_i = 1]$.

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

Summary

Variance

- Variance: $var[X] := E[(X - E[X])^2] = E[X^2] - E[X]^2$
- Fact: $var[aX + b] = a^2 var[X]$
- Thm.: If X, Y are indep., $Var(X+Y) = Var(X) + Var(Y)$.
- $U[1,\ldots,n] : Pr[X = m] = \frac{1}{n}, m = 1,\ldots,n$;

  $E[X] = \frac{n+1}{2}$; $var(X) = \frac{n^2-1}{12}$.;

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

  $E[X] = np$; $var(X) = \; = np(1-p)$.

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

  $E[X] = \frac{1}{p}$; $var[X] = \frac{1-p}{p^2}$.