Prob. Dist^n: Formal name for Prob. assignment for all values in D_X.

(X discrete $\iff D_X$ is finite / countably infinite)

Ex: 1) If the coin is fair,

$P(X=0) = P(\{T\}) = 0.5$

$P(X=1) = P(\{H\}) = 0.5$

2) If the die is fair,

$P(X=1) = P(X=2) = \dots = P(X=6) = \frac{1}{6} \longrightarrow$ Discrete Uniform dist^n (N=6)

Defn: Probability mass function (pmf): $p(x) := p_X(x) = P(X=x) \quad \forall x \in \mathbb{R}$

By defn, $p_X(x) = 0 \quad \forall x \notin D_X$.

Requirements: 1. $p_X(x) \geq 0 \quad \forall x$

2. $\sum_{x \in D_X} p_X(x) = P(\mathcal{S}) = 1$ (The events $\{X=x\}$ for all possible $x \in D_X$ are disjoint and exhaustive)

$1, 2 \Rightarrow p_X(x) \leq 1 \quad \forall x$

Eg. 2. $Y = \begin{cases} 1 & \text{if } X \geq 4 \\ 0 & \text{if } X \leq 3 \end{cases}$ (Fair Die)

$p_Y(0) = p_X(1) + p_X(2) + p_X(3)$ $(= P(X \in \{1,2,3\}) = P(X \leq 3))$

$= \frac{3}{6} = 0.5$

$p_Y(1) = 0.5$

$Z = \begin{cases} 1 & \text{if } X = 6 \\ 0 & \text{if } X < 6 \end{cases}$ $\quad p_Z(0) = 5/6$, $\quad p_Z(1) = 1/6$

Y, Z are both Bernoulli.

In general, for a Bernoulli RV X, we can write:

$$\left.\begin{aligned} P(X=0) &= p_X(0) = \alpha \\ P(X=1) &= p_X(1) = 1-\alpha \end{aligned}\right\} \text{ for any value } 0 < \alpha < 1.$$

α: Parameter (of a family of dist^{ns})

$$P(X=x) = p(x;\alpha) = p(x|\alpha)$$

Prob. dist^n: all prob. assignments for all x in D_X.

Prob. dist^n family: all prob. dist^{ns} by varying α.

In a dist^n family, all dist^{ns} are of the same form. Only the parameter value is changing.

e.g. 1. In eg. 2 both Y & Z are Bernoulli but with $\alpha = 0.5$ and $\alpha = 5/6$ respectively.

2. In eg. 2, X is discrete uniform with N=6.

In general, for Discrete Uniform dist^n with N.

$D_X = \{1, 2, \dots, N\}$

$p_X(x|N) = \frac{1}{N} \quad \forall x \in D_X$.

Ex:- Starting at some time T_0, we observe the gender of each new born baby until a boy (B) is born.

X = # of babies (up to & including the 1st boy)

Assume successive births are indep.

Let $P(B) = p = 1 - P(G)$ $(0 < p < 1)$

$\mathcal{S} = \{B, GB, GGB, \dots\}$

$D_X = \{1, 2, 3, 4, \dots\}$

$p_X(1) = P(X=1) = P(B) = p$ $\quad P(G)P(B) = (1-p)p$

$p_X(2) = P(X=2) = P(GB)$

$p_X(3) = P(X=3) = P(GGB) = (1-p)^2 p$

$\vdots$

$$p_X(x;p) = p_X(x) = \begin{cases} (1-p)^{x-1}\, p, & \text{if } x \in D_X \\ 0, & \text{if } x \notin D_X \end{cases}$$ Geometric dist^n with parameter p.

+ve Integer

Disjoint

$$P(X \le n) = P\left(\{X=1\} \cup \{X=2\} \cup \dots \cup \{X=n\}\right)$$
$$= \sum_{x=1}^{n} P(X=x) = P(X=1) + \dots + P(X=n)$$
$$= \sum_{x=1}^{n} p_X(x).$$

<u>Def^n</u>: Cumulative dist^n function (cdf): $F_X(x) \overset{\text{def}^n}{=} P(X \le x)$

$$= \sum_{y:\, y \le x} p_X(y) \quad (\text{When } X \text{ is discrete})$$

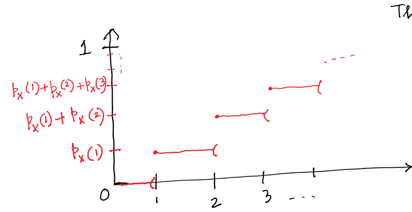
So, $$F_X(n) = \sum_{x=1}^{n} (1-p)^{x-1} \cdot p$$
$$= p + p(1-p) + \dots + p(1-p)^{n-1} \quad : \text{Geometric Series}$$
$$= p\cdot\left[1 + (1-p) + \dots + (1-p)^{n-1}\right]$$
$$= p\left[\frac{(1-p)^n - 1}{(1-p) - 1}\right] = 1 - (1-p)^n \longrightarrow 1 \text{ as } n \to \infty$$
(as $0 < p < 1$)

If $p = 0.51$, $F_X(10) \approx 1$.

$F_X(2.5) = p_X(1) + p_X(2) = F_X(2)$.

In general, $F_X(x) = F_X(\lfloor x \rfloor)$ (eg. $\lfloor 2.5 \rfloor = 2$, $\lfloor -1.2 \rfloor = -2$)

The largest integer $\le x$.



$F_X(x) = P(X \le x) = 0$ for any $x < 1$ as X must be ≥ 1.

$F_X(1) = P(X \le 1) = P(X=1) = p_X(1)$

$F_X(x) = p_X(1)$ for any $x \in (1,2)$

$F_X(2) = P(X \le 2) = P(X=1) + P(X=2) = p_X(1) + p_X(2)$

$F_X(x) = p_X(1) + p_X(2)$ for any $x \in (2,3)$

Continue this way.

Lecture 13

Thursday, June 29, 2017 12:58 PM

Support of discrete X : $D_X = \{x : p_X(x) > 0\}$

CDF : $F(x) = P(X \le x) = \sum_{y : y \le x} p(y)$

If $X \sim Geo(p)$, $F(n) = 1 - (1-p)^n$ for all $n = 1, 2, \dots$

$F(0) = 0$

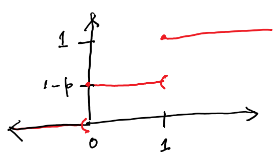
$F(n) \uparrow 1$ as $n \uparrow \infty$ $\left[\sum_{x=1}^{\infty} p(x) = \sum_{x=1}^{\infty} p(1-p)^{x-1} = \frac{p}{1-(1-p)} = 1 \quad \left(\begin{matrix} a = p \\ r = 1-p \end{matrix}\right)\right]$

$F(x) = F([x]) \ \forall x \in \mathbb{R}$

e.g. $F(2.5) = P(X \le 2.5) = P(X=1) + P(X=2)$
$= p(1) + p(2) = F(2)$ $([2.5] = 2)$

$F(2.9999) = F(2) = p(1) + p(2)$

If $\underset{\substack{\uparrow \\ \text{\# Success}}}{X} \sim Ber(\underset{\substack{\uparrow \\ \text{Prob. of Success}}}{p})$, $p(0) = 1-p$, $p(1) = p \Rightarrow p(x) = p^x (1-p)^{1-x}$ for $x = 0, 1$



If $x < 0$, $F(x) = 0$

$F(0) = p(0) = 1-p$

$F(x) = 1-p$, $\forall x \in (0,1)$

$F(1) = p(0) + p(1) = 1$

$F(x) = 1$, $\forall x > 1$

$$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1-p, & \text{if } 0 \le x < 1 \\ 1, & \text{if } x \ge 1 \end{cases}$$

General Results:

- $F(x) \uparrow 1$ as $x \uparrow \infty$, $F(\infty) = P(X \le \infty) = 1$
- $F(x) \downarrow 0$ as $x \downarrow -\infty$, $F(-\infty) = P(X \le -\infty) = 0$
- F is always non-decreasing. [If $x > y$, $\{X \le y\} \subseteq \{X \le x\}$
  $\Rightarrow P(X \le y) \le P(X \le x)$
  $\Rightarrow F(y) \le F(x)$]
- If X is discrete, F is a step fn.
- F is always right continuous.

$X \sim Geo(p) \longrightarrow$ ① Construct CDF from PMF
② Reconstruct PMF from CDF.

① $F(0) = 0$
$F(1) = p(1)$
$F(2) = p(1) + p(2)$
$F(3) = p(1) + p(2) + p(3)$
$\vdots$
For non-integer x, $F(x) = F([x])$

This is true if $D_X = \{1, 2, 3, \dots\}$

② $p(1) = F(1) - F(0) = F(1)$
$p(2) = F(2) - F(1)$
$p(3) = F(3) - F(2)$
$\vdots$
$F(3) - F(1) = p(2) + p(3)$
$= P(2 \le X \le 3)$
$F(4) - F(2) = P(3 \le X \le 4)$
In general, $P(a \le X \le b) = F(b) - F(a-1)$ (for +ve integers a & b)
$\quad\quad\quad\quad \overset{\shortparallel}{P(a-1 < X \le b)}$
Put $a = b$, $P(X = a) = F(a) - F(a-1)$
$\quad\quad\quad \overset{\shortparallel}{p(a)}$

If $D_X = \{0, 0.5, 1, 1.5, \dots\}$

$p(1) = \underbrace{F(1)}_{p(0) + p(0.5) + p(1)} - \underbrace{F(0.5)}_{p(0) + p(0.5)}$

Eg. X = # courses taken by a random student

$D_X = \{1, 2, \dots, 7\}$

Eg. X

$D_X = \{1, 2, \dots, 7\}$

X	1	2	3	4	5	6	7	Total
# of Students	150	450	1950	3750	5850	2550	300	15000

Avg. # of Courses (per student) (Population Mean)

$$= \frac{1\times150 + 2\times450 + \dots + 7\times300}{15000}$$

$$= 1\times\left(\frac{150}{15000}\right) + 2\times\left(\frac{450}{15000}\right) + \dots + 7\times\left(\frac{300}{15000}\right)$$

$$= 1\times p_X(1) + 2\times p_X(2) + \dots + 7\times p_X(7)$$

$$= 4.57$$

Defn: Popl^n Mean / Popl^n avg. / Expectation / Expected Value (of X)

$$\mu := \mu_X \overset{\text{defn}}{=} \sum_{x\in D_X} x\, p_X(x)$$

$(= E(X))$

If $X \sim \text{Ber}(p)$, $\mu = 0\cdot p(0) + 1\cdot p(1)$ (By defn.)

$$= 0\cdot(1-p) + 1\cdot p = p$$

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

w.p. with prob.

Eg. (contd.) Each course requires 2 books.
Y = # books a random student needs

Y	2	4	6	8	...	14
# of Students	150	450	1950	3750	...	300

$D_Y = \{2, 4, \dots, 14\}$

$$\mu_Y = \sum_{y\in D_Y} y\, p_Y(y) = 2\times\frac{150}{15000} + 4\times\frac{450}{15000} + \dots + 14\times\frac{300}{15000}$$

$$= 2\left(1\cdot p_X(1) + 2\cdot p_X(2) + \dots + 7\cdot p_X(7)\right)$$

$$= 2\cdot\mu_X = 2\times4.57 = 9.14$$

$$EY = \sum_{y\in D_Y} y\cdot p_Y(y)$$

$$= \sum_{x\in D_X} 2x\cdot p_X(x) \quad \left[\text{Put } y = 2x\right]$$

We have $p_Y(y) = p_X(x)$ if $y = 2x$ $\forall x = 1, 2, \dots, 7$

$$Y = 2X \Rightarrow \mu_Y = 2\mu_X$$

If $h(\cdot)$ is a fn. of X, then

$$E[h(X)] = \mu_{h(X)} \overset{\text{defn.}}{:=} \sum_{x\in D_X} h(x)\, p_X(x).$$

We have seen: $\mu_{2X} = 2\mu_X$ (With the function $h(x) = 2x$)

Lecture 14

Friday, June 30, 2017 1:00 PM

If $Y = h(X)$,

$$\mu_Y = E(Y) = \sum_{y \in D_Y} y\, p_Y(y)$$
$$= \sum_{x \in D_X} h(x)\, p_X(x)$$

We have seen this for $h(x) = 2x$.

X	x_1	x_2	....	x_n
$p_X(x)$	$p_1 = p_X(x_1)$	$p_2 = p_X(x_2)$	....	$p_n = p_X(x_n)$
$h(X)$	$h(x_1)$	$h(x_2)$		$h(x_n)$

Then, $\mu_X = \sum_{x \in D_X} x\, p_X(x) = \sum_{i=1}^{n} x_i p_i$

and $\mu_{h(X)} = E(h(X)) = \sum_{x \in D_X} h(x)\, p_X(x) = \sum_{i=1}^{n} h(x_i)\, p_i$

Moreover, we have seen that for $h(x) = 2x$,

$\mu_{h(X)} = 2\mu_X$, i.e. $\boxed{E[h(X)] = h(E[X])}$

This is NOT true in general, unless $h(x) = ax + b$ (for constants a, b), i.e. $h(\cdot)$ is a linear fn.

Proof: $\mu_{h(X)} = E[h(X)] = \sum_{x \in D} (ax+b)\, p(x)$

$$= a \sum_{x \in D} x\, p(x) + b \sum_{x \in D} p(x)$$
$$= a\, E(X) + b.1 = a\mu_X + b = h(\mu_X) = h(E[X])$$

i.e. $\boxed{\mu_{aX+b} = a\mu_X + b \quad \text{OR, } E(aX+b) = aE(X)+b}$ → Put $a=1$ → $\mu_{X+b} = \mu_X + b$; Put $b = 0$ → $\mu_{aX} = a\mu_X$

- You "expect" the sample average to get close to the population average in the long run.

If you observe $X_1, X_2, \ldots, X_n$ from n indep. trials,

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow E(X) = \mu_X \text{ as } n \to \infty$$
$$\frac{1}{n}\sum_{i=1}^{n} h(X_i) \longrightarrow E[h(X)] = \mu_{h(X)} \text{ as } n \to \infty$$

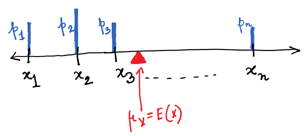
(Law of large numbers)

- Physical interpretation:

Put mass $p_i = p_X(x_i)$ at all the x_i's on the horizontal axis.

$\mu_X = E(X)$ is the "center of mass" or "fulcrum" which the axis will balance horizontally.

It is a measure of "location" or "central tendency"

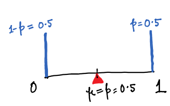


$\boxed{\mu_{X+b} = \mu_X + b}$ is easy to see from this interpretation of μ. If you move all the masses by a distance b (in the same direction), then to balance the axis the fulcrum needs to move by the same distance in that direction!

Eg. 1. $X \sim Ber(p)$

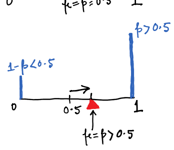
We know: $\mu = p$

If $p = 0.5$,



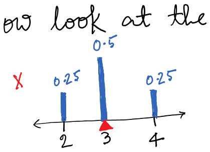
Two equal masses at 0 and 1.
→ Balanced at $\mu = 0.5$.

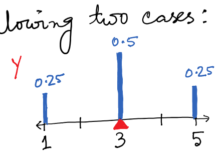
If $p > 0.5$,
$1 - p < 0.5$



Mass at 1 is more
→ Fulcrum should move to the right side.

2. Now look at the following two cases:





$\mu_X = 2 \times .25 + 3 \times .5 + 4 \times .25 = 3$

$\mu_Y = 1 \times .25 + 3 \times .5 + 5 \times .25 = 3$

Clearly, population mean μ cannot distinguish between X and Y.

We need measures of 'spread'.

- Variance of X:

Sample Variance (of X): $\frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2$: Sample Mean of Squared deviation of X's from the sample mean of X's.

the sample mean of ^s.

Extend this to:

Population Variance (of X) $:= Var(X) := V(X) := \sigma_X^2$

$$\overset{\text{def}^n}{=} E\left[\left(X-\mu_X\right)^2\right]$$: Popl^n Mean of squared deviation of X from the popl^n mean of X.

This def^n works for both discrete and conts. cases.

Standard deviation (of X) $:= SD(X) := \sigma_X \overset{\text{def}^n}{=} +\sqrt{Var(X)}$

With $h(x) = (x-\mu_X)^2$, we have:

X	x_1	x_2	---	x_n
$p_X(x)$	$p_1 = p_X(x_1)$	$p_2 = p_X(x_2)$	---	$p_n = p_X(x_n)$
$h(x)$	$(x_1-\mu_X)^2$	$(x_2-\mu_X)^2$	---	$(x_n-\mu_X)^2$

Here $\mu_X = \sum_{i=1}^{n} x_i p_i$ is just a constant.

Hence, $V(X) = E[X-\mu_X]^2 = E[h(X)]$
$$= \sum_{i=1}^{n} (x_i-\mu_X)^2 p_i \text{, where } \mu_X = \sum_{i=1}^{n} x_i p_i$$

Eg. 1 (contd.) For $X \sim Ber(p)$

We have: $\mu = p$

$$\begin{aligned} \text{So } V(X) &= E(X-\mu)^2 = E(X-p)^2 \\ &= (0-p)^2 p(0) + (1-p)^2 p(1) \\ &= p^2 \cdot (1-p) + (1-p)^2 p \\ &= p(1-p)(\cancel{p}+1-\cancel{p}) \\ &= p(1-p) \end{aligned}$$

Eg. 2. (contd.)

Check that $V(X) < V(Y)$ in Eg. 2 above.

Clearly Y is more 'spread' compared to Y.

- Properties of Variance:

1. Shortcut formula: $\boxed{V(X) = E(X^2) - E^2(X)}$ $\Leftrightarrow$ $\boxed{\sigma_X = \sqrt{E(X^2) - E^2(X)}}$

Proof:
$$\begin{aligned} V(X) &= E(X-\mu)^2 \\ &= \sum_{x\in D} (x-\mu)^2 p(x) \\ &= \sum_{x\in D} (x^2 - 2x\mu + \mu^2) p(x) \\ &= \sum_{x\in D} x^2 p(x) - 2\mu \sum_{x\in D} x p(x) + \mu^2 \sum_{x\in D} p(x) \\ &= E(X^2) - 2\mu\cdot\mu + \mu^2 \cdot 1 \\ &= E(X^2) - \mu^2 = E(X^2) - E^2(X) \end{aligned}$$

Eg. 1. (contd.) $X \sim Ber(p)$

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

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

$$\Rightarrow X^m = \begin{cases} 0 & \text{w.p. } 1-p \\ 1 & \text{w.p. } p \end{cases} \quad (\text{for any integer } m \geq 1)$$

$\Rightarrow X^m \sim Ber(p), \forall m \geq 1$ (Eg. $X^{79} \sim Ber(p) \Rightarrow E(X^{79}) = p$)

In particular, $X^2 \sim Ber(p)$

$\Rightarrow E(X^2) = p$ [Can also get from: $E(X^2) = \sum_{x=0}^{1} x^2 p(x)$
$= 0^2 \cdot p(0) + 1^2 \cdot p(1)$
$= 0 \cdot (1-p) + 1 \cdot p = p$]

Now, $V(X) = E(X^2) - E^2(X)$
$= p - (p)^2$
$= p - p^2 = p(1-p)$

which we have already seen.

2. $V(X) \geq 0$

as $V(X) = \sum_{x \in D} \underbrace{(x-\mu)^2}_{\geq 0} \underbrace{p(x)}_{\geq 0}$ is a sum of non-negative numbers.

This shows that $E(X^2) - E^2(X) \geq 0$
i.e. $E(X^2) \geq \{E(X)\}^2$
i.e. $E[h(X)] \geq h(E[X])$ if $h(x) = x^2$.
Since $h(x) = x^2$ is NOT a linear fn., the equality $E[h(X)] = h(E[X])$ does not hold.

3. If $X = c$ w.p. 1, $V(X) = 0$.
↑
Some constant

Proof: $E(X^2) = c^2 \cdot 1 = c^2$
and $E(X) = c \cdot 1 = c$
$\Rightarrow V(X) = c^2 - (c)^2 = c^2 - c^2 = 0.$

This is called a 'degenerate' RV.

In fact, if $V(X) = 0$, X must be a degenerate RV.

4. If a, b are constants, $V(aX+b) = a^2 \cdot V(X)$

Defn: $Var(h(X)) := V(h(X))$
$\overset{\text{defn}}{=} E\Big[\underbrace{h(X) - \mu_{h(X)}}_{\text{Treat this as another fn. } K(X).}\Big]^2$: Popln mean of squared deviation of $h(X)$ from the popln mean of $h(X)$.

$= \sum_{x \in D} \left[h(x) - \mu_{h(X)}\right]^2 p(x)$

Proof: $V(aX+b) = \sum_{x \in D} (ax+b-\mu_{aX+b})^2 p(x)$
$= \sum_{x \in D} (ax+\cancel{b} - a\mu_X - \cancel{b})^2 p(x)$
$= a^2 \sum_{x \in D} (x-\mu_X)^2 p(x)$
$= a^2 V(X).$

It also shows that $V(aX+b)$ does NOT depend on b.
As a special case, $V(X+b)$ does NOT depend on b.
If all the values are shifted by an amount b, but the probabilities remain unchanged, then Variance, being a measure of spread, should NOT change!

i.e. $\boxed{\begin{aligned}\sigma^2_{aX+b} &= a^2\sigma_X^2 \\ \sigma_{aX+b} &= |a|\sigma_X\end{aligned}}$ Hence

Special case: $\sigma_{aX} = |a|\sigma_X$ (Put $b=0$) and $\sigma_{X+b} = \sigma_X$ (Put $a=1$)

This works even when $a=0$.
If $a=0$, $aX+b = b$ w.p. 1 (i.e. degenerate RV with value b).
Hence, $V(aX+b) = 0 = a^2 \cdot V(X)$ [as $a=0$]
$\Rightarrow \sigma_{0.X+b} = 0.\sigma_X = 0$

- Section 3.4:

Binomial Experiment — i) Consists of n trials (n: integer ≥ 1)
ii) Each trial has $\underbrace{\text{2 possible outcomes}}_{\text{Dichotomous}}$ (S & F)
iii) Trials are indep.
iv) P(S) and P(F) don't change across trials.

Eg. 1. n Flips of a coin
$S = \{H\}$, $F = \{T\}$.
2. n Rolls of a die.
$S = \{1,2,3\}$
$F = \{4,5,6\}$

Discrete RV: $X = \#$ of S's in n trials $\sim \underbrace{Bin(n,p)}_{\substack{\text{a 2-parameter family} \\ \text{with parameters } n \text{ and } p. \\ (n:\text{ integer} \geq 1 \text{ and } 0 \leq p \leq 1)}}$ [with $p = P(S) = 1-P(F)$]

Support: $D_X = \{0, 1, 2, \dots, n\}$

PMF: $p_X(x) := P(X=x)$

$\overset{\text{Notation}}{=} b(x;n,p)$

$= \left(\begin{array}{l}\#\text{ ways that you can} \\ \text{have } x \text{ S's out of} \\ n \text{ trials}\end{array}\right) \times \left(\begin{array}{l}\text{Probability of each} \\ \text{of those ways}\end{array}\right)$

$$= \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & \text{if } x = 0,1,2,\dots,n \\ 0, & \text{otherwise} \end{cases}$$

(Eg. $P(X=0) = b(0;n,p)$
$= \binom{n}{0} p^0 (1-p)^{n-0}$
$= (1-p)^n$
$P(X=1) = b(1;n,p)$
$= \binom{n}{1} p^1 (1-p)^{n-1}$
$= np(1-p)^{n-1}$ etc.)

CDF: $F_X(x) := P(X \leq x)$

$\overset{\text{Not}^n}{=} B(x;n,p)$

$$= \begin{cases} \sum_{y=0}^{[x]} b(y;n,p), & \text{if } 0 \leq x \leq n. \\ 0, & \text{if } x < 0. \\ 1, & \text{if } x > n. \end{cases}$$

Moreover, $E(X) = np$ and $V(X) = np(1-p)$.

- Section 3.6:

Poisson RV — does NOT arise exactly from a specific random expt.
$X \sim Poi(\lambda)$: 1-parameter family with parameter $\lambda > 0$.
$D_X = \{0, 1, 2, \dots\}$

PMF: $P(X=x) = p(x;\lambda)$

$$= \begin{cases} \dfrac{e^{-\lambda}.\lambda^x}{x!}, & \text{for } x = 0,1,2,\dots \\ 0, & \text{otherwise} \end{cases}$$

(Eg. $P(X=0) = p(0;\lambda)$
$= \dfrac{e^{-\lambda}.\lambda^0}{0!} = e^{-\lambda}$
$P(X=1) = p(1;\lambda)$
$= \dfrac{e^{-\lambda}.\lambda^1}{1!} = \lambda e^{-\lambda}$ etc.)

CDF: $$P(X \leq x) = \begin{cases} \sum_{y=0}^{[x]} \dfrac{e^{-\lambda}.\lambda^y}{y!}, & \text{for } y \geq 0. \\ 0, & \text{for } y < 0. \end{cases}$$

Moreover, $E(X) = \lambda = V(X)$.

We will discuss these 2 families in detail in the next class.

Lecture 15

Wednesday, July 5, 2017 12:58 PM

Section 3.4:

Binomial Experiment eg. 1) n Flips of a coin $\to S=\{H\}, F=\{T\}$

2) n Rolls of a die $\to S=\{2,4,6\}, F=\{1,3,5\}$

3) Balls marked 'S' & 'F' in a box
– Draw 1 ball at a time (n times) with replacement

4) $N=50$ restaurants in an area
$\to 15$ violate health code, 35 don't
$n=5$ restaurants are chosen. (without replacement)
S = No violation

$$P(S_1)=\frac{35}{50}, \quad P(S_2)=P(S_2\cap S_1)+P(S_2\cap F_1)$$
$$=P(S_2|S_1)P(S_1)+P(S_2|F_1)P(F_1)$$
$$=\frac{34}{49}\times\frac{35}{50}+\frac{35}{49}\times\frac{15}{50}$$
$$=\frac{35}{49\times 50}(34+15)=\frac{35}{50}$$

$$P(S_1\cap S_2)=\frac{34\times 35}{49\times 50}\neq\left(\frac{35}{50}\right)^2=P(S_1)\cdot P(S_2) : \text{NOT indep.}$$

Replace 35, 15 and 50 by 35000, 15000 and 50000.

$$P(S_1\cap S_2)=\frac{35000\times 34999}{50000\times 49999}\approx\left(\frac{35000}{50000}\right)^2=P(S_1)P(S_2)$$

Near indep. holds for $n=5$.

- If Population Size N is too large compared to sample size n, i.e. $N\geqslant 20n$, then Without replacement sampling will be an approximate Binomial Experiment.

Here, $N=50$, $n=5$ originally $\to \frac{N}{n}=10<20$ (NOT binomial)
but later, $N/n=10000\geqslant 20 \to$ (Approx. binomial)

Now, $X=$ # S's in n trials
Then, $X\sim Bin(n,p)$ $[p=P(S)=1-P(F)]$
$D=\{0,1,\dots,n\}$
If $n=1$, $X\sim Ber(p)$

Parameters: $n\geqslant 1$: integer
$0\le p\le 1$
If $p=0$, $X=0$ w.p. 1
If $p=1$, $X=n$ w.p. 1
} X is degenerate

$$X\sim Bin(n,p)\Rightarrow X=Y_1+\dots+Y_n \quad (Y_i=\#\text{ S's in } i^{th}\text{ trial} \overset{iid}{\sim} Ber(p))$$

iid: indep. & identically distributed

Eg. $n=3\to(Y_1,Y_2,Y_3)\overset{iid}{\sim}Ber(p)$
$X=Y_1+Y_2+Y_3$
$|\mathcal{S}|=2^3=8$
$\mathcal{S}=\{000,001,\dots,111\}$

$s\in\mathcal{S}$	$X(s)$	$P(\{s\})$
000	0	$(1-p)^3$
001	1	$p(1-p)^2$
010	1	$p(1-p)^2$
100	1	$p(1-p)^2$
011	2	$p^2(1-p)$
101	2	$p^2(1-p)$
110	2	$p^2(1-p)$
111	3	p^3

Check that these add up to 1!

$P(\{0\})=1-p=1-P(\{1\})$

$$P(X=0)=P(\{000\})=(1-p)^3$$
$$P(\{001\}\cup\{010\}\cup\{100\})=3p(1-p)^2$$

$P(X=1) = P(\ldots)$

$P(X=2) = P(\{011\} \cup \{101\} \cup \{110\}) = 3p^2(1-p)$

$P(X=3) = P(\{111\}) = p^3$

In general,

$b(x;n,p) = P(X=x) = P(\text{There are } x \text{ S's in } n \text{ trials})$

Notation

$= \left(\begin{array}{c}\#\text{ outcomes that}\\ \text{have } x \text{ S's}\end{array}\right) \times \left(\begin{array}{c}\text{Prob. of each}\\ \text{outcome with}\\ x \text{ S's}\end{array}\right)$ [$\because$ Equally likely]

$= \binom{n}{x} p^x (1-p)^{n-x}$ [$\forall x \in \{0,1,\ldots,n\}$]

For binomial PMF & CDF, we can use the tables also. $(n = 5, 10, \ldots, 25)$

- $X = Y_1 + Y_2 + \cdots + Y_n$

$E(Y_i) = E(\#\text{ S's in } i^{th} \text{ trial}) = p$ (as $Y_i \sim Ber(p)$)

$E(X) = E(\#\text{ S's in } n \text{ trials}) = np$

Actually, $E(X) = E(Y_1) + \cdots + E(Y_n) = np$

Also, $V(Y_i) = p(1-p)$

We'll see later: $V(X) = V(Y_1) + \cdots + V(Y_n)$ [By indep.]

$= np(1-p)$

$\Rightarrow \sigma_X = \sqrt{np(1-p)} = \sqrt{npq}$ $(q = 1-p)$

Skip 3.5. Hypergeometric $\rightarrow$ Sampling without replacement.

Neg. Binomial $\rightarrow$ Waiting for r^{th} Success in Bernoulli trials. $(r \geq 1)$

Section 3.6: Poisson Dist^n — Useful for analyzing count/number of certain events.

$X \sim Poi(\lambda)$ (parameter: λ)

$D = \{0, 1, 2, \ldots\}$

A Prob. dist^n on an infinite support requires an infinite series of positive terms with finite sum.

Eg. Geometric series: $a + ar + ar^2 + \cdots = \frac{a}{1-r}$ (with $0 < r < 1$)

$\Rightarrow \left(\frac{a}{1-r}\right)^{-1} a + \left(\frac{a}{1-r}\right)^{-1} ar + \cdots = 1$

$\Rightarrow \underbrace{(1-r)} + \underbrace{r(1-r)} + \underbrace{r^2(1-r)} + \cdots = 1$

Set $P(X=1)$, $P(X=2)$, $P(X=3)$, ... etc. [and put $p = 1-r$]

We get the Geometric dist^n PMF.

Now, look at the exponential series: $e^\lambda = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots$

$\Rightarrow 1 = \underbrace{e^{-\lambda}} + \underbrace{\lambda e^{-\lambda}} + \underbrace{\frac{\lambda^2}{2!} e^{-\lambda}} + \cdots$

Set $P(X=0)$, $P(X=1)$, $P(X=2)$, ... etc.

and you get the Poisson(λ) dist^n PMF.

Poisson PMF: $p(x;\lambda) = \begin{cases} \dfrac{e^{-\lambda}\cdot\lambda^x}{x!} & \text{for } x = 0,1,2,\ldots \\ 0 & \text{o.w.} \end{cases}$

Eg. X = # Animals in a trap in a day.
Given, $X \sim Poi(4.5)$ [Rate = 4.5]

$$P(X=5) = p(5;4.5) = \frac{e^{-4.5}\cdot(4.5)^5}{5!} = 0.1708$$

$$P(X \le 5) = P(X=0) + P(X=1) + \cdots + P(X=5)$$
$$= \sum_{x=0}^{5} \frac{e^{-4.5}\cdot(4.5)^x}{x!}$$

For x-values $\ll \lambda$ or $\gg \lambda$, $p(x;\lambda) \approx 0$. [Similarly, for $x \ll np$ or $x \gg np$, $b(x;n,p) \approx 0$]

Prop: If $X \sim Bin(n,p)$ and if $n \to \infty$ and $np \to \lambda$,

$$P(X=x) = b(x;n,p) \longrightarrow p(x;\lambda) = \frac{e^{-\lambda}\cdot\lambda^x}{x!}$$
$$(\forall x = 0,1,2,\ldots)$$

- Rules of Thumb:

Use Poisson approx. if $n > 50$ and $np < 5$.
Or alternatively, if $n \geq 100$, $p \le 0.01$ and $np \le 20$.

eg. $n = 200$ items sampled, each with defective prob. $p = 0.005$.
Prob. of at most 1 defective item:

$$P(X \le 1) = P(X=0) + P(X=1) = \binom{200}{0}(.005)^0(0.995)^{200} + \binom{200}{1}(.005)^1(0.995)^{199}$$
$$= 0.7357597$$

With Poisson approx. ($\lambda = np = 1$)

$$P(X \le 1) = P(X=0) + P(X=1) = \frac{e^{-1}\cdot 1^0}{0!} + \frac{e^{-1}\cdot 1^1}{1!} = 0.7357589$$

- Mean and Variance: Let $Y \sim Poi(\lambda)$.

Let $X_n \sim Bin(n,p)$ and $n \to \infty$ with $np \to \lambda$. (clearly, $p \to 0$)

$\therefore P(X_n = x) \longrightarrow P(Y=x)$ $\forall x = 0,1,2,\ldots$ (as $n \to \infty$)

$\Rightarrow E(X_n) \longrightarrow E(Y)$ and $V(X_n) \longrightarrow V(Y)$. ⊛

Now, $E(X_n) = np$ and $V(X_n) = np(1-p)$.

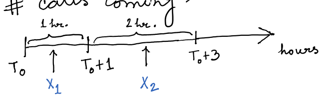
As $np \to \lambda$ and $p \to 0$ as $n \to \infty$,
$E(X_n) \to \lambda$ and $V(X_n) \to \lambda\cdot(1-0) = \lambda$ as $n \to \infty$. ⊛⊛

From ⊛ and ⊛⊛,
$E(Y) = \lambda$ and $V(Y) = \lambda$.
Both mean and variance of $Poi(\lambda)$ are λ.

- Poisson Process: Used to model random events distributed in time or space.

e.g. Arrival of phone calls at an exchange or customers at a store.
It has two properties:
(i) # events in any finite interval has a Poisson dist^n.
(ii) # events in disjoint intervals are indep. RV's.
Furthermore, if the rate parameter λ remains fixed over time, the dist^n (equivalently, the mean) of the # events in any interval depends on the length of the interval. [Mean # events in $(a,b] = \lambda(b-a)$]

eg. # calls coming to an exchange observed from time T_0 (say, 6:00 AM). Let the rate λ be 5 calls/hour.



X_1 = # calls between 6 AM and 7 AM.
X_2 = # calls between 7 AM and 9 AM.
Now, $X_1 \sim Poi(5)$, $X_2 \sim Poi(10)$ and X_1 is indep. of X_2.
So, P(exactly 2 calls between 6 and 7 AM and 3 calls between 7 and 9 AM) $= P(X_1 = 2, X_2 = 3)$
$= P(X_1 = 2)\cdot P(X_2 = 3)$
$= \dfrac{e^{-5}(5)^2}{2!}\cdot\dfrac{e^{-10}(10)^3}{3!}$

Now, if X_3 = # calls between 6:30 AM and 7:30 AM.
Both X_1 and X_3 have $Poi(5)$ dist^n, but they are NOT indep. (Overlapping intervals)

- Review: If a fair 6-sided die is rolled 10 times, what is the probability that there is exactly 3 '1's and exactly 2 '2's?

Let us club 3-6 all together as ⊛, so we have 3 basic outcomes from a single roll: '1', '2' and ⊛, so that $P(\{1\}) = P(\{2\}) = \frac{1}{6}$ and $P(\{⊛\}) = \frac{4}{6}$.

… 3 '1's, 2 '2's and 5 ⊛'s can be arranged) × (Prob. of each arrangement)

- Exactly 2 2's.

Let us club 3–6 all together as ⊛, so we have 3 basic outcomes from ...

Probability = (# ways 3 '1's, 2 '2's and 5 ⊛s can be arranged) × (Prob. of each arrangement)

$$= \underbrace{\left[\binom{10}{3}\binom{7}{2}\binom{5}{5}\right]} \times \underbrace{\left[\left(\frac{1}{6}\right)^3\left(\frac{1}{6}\right)^2\left(\frac{4}{6}\right)^5\right]}_{\text{By indep. of successive rolls.}} = \underbrace{\frac{10!}{3!\,2!\,5!}}_{\text{This number can be derived in another way.}} \times \left(\frac{1}{6}\right)^3\left(\frac{1}{6}\right)^2\left(\frac{4}{6}\right)^5.$$

ways to choose 3 places for the '1's out of 10 places

ways to choose 2 places for the '2's out of the remaining 7 places.

ways to choose 5 places for the ⊛s out of the remaining 5 places. (ONLY 1 way to do it)

$$\frac{10!}{3!\,2!\,5!}$$

10! → # ways 10 distinct items can be arranged

3! — # arrangements you lose as the 3 '1's are NOT unique for each arrangement of other items.

2! 5! — Similar terms as 2 '2's are NOT unique and 5 ⊛s are NOT unique.

- $E[f(X)+g(X)] = E[f(X)] + E[g(X)]$

Proof: $LHS = \sum_{x\in D}[f(x)+g(x)]\,p_X(x)$ [Take $h(x)=f(x)+g(x)$]

$= \sum_{x\in D} f(x)p_X(x) + \sum_{x\in D} g(x)p_X(x)$

$= E[f(X)] + E[g(X)]$ (if at least one of $E[f(X)]$ and $E[g(X)]$ is finite in absolute value)

eg: $\left(1+\frac{1}{2}+\frac{1}{3}+\cdots\right) = \infty$

$\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots\right) = c < \infty$

Put $P(X=x) = \frac{1}{cx^2}$ $[x=1,2,\dots]$

$D = \{1,2,\dots\}$

This is a valid PMF. ($p(x)=\frac{1}{cx^2}>0\ \forall x=1,2,\dots$ and $\sum_{x=1}^{\infty}p(x) = \frac{1}{c}\cdot\sum_{x=1}^{\infty}\frac{1}{x^2} = \frac{1}{c}\cdot c = 1$)

$E(X) = \sum_{x=1}^{\infty} x\cdot p(x) = \sum_{x=1}^{\infty} x\cdot\frac{1}{cx^2} = \frac{1}{c}\sum_{x=1}^{\infty}\frac{1}{x} = \infty$

Now, take $f(x)=x$, $g(x)=-x$

$f(x)+g(x)\equiv 0$

$\Rightarrow E[f(X)+g(X)] = 0$

But, $E[f(X)] = E(X) = +\infty$

and $E[g(X)] = E(-X) = -E(X) = -\infty$

So, $E[f(X)+g(X)] = E[f(X)] + E[g(X)]$ does NOT hold.

- Exam Problem #6: 20 MCQ's, 5 options for each.

S = Correct answer for a question

$p = P(S) = \frac{1}{5} = 0.2$

X = # Correct answers in the Exam

$X \sim Bin(20, 0.2)$

(a) $P(X=5 \mid X\geq 1) = \dfrac{P(\{X=5\}\cap\{X\geq 1\})}{P(X\geq 1)} = \dfrac{P(X=5)}{1-P(X=0)} = \dfrac{b(5;20,0.2)}{1-b(0;20,0.2)}$

$= \dfrac{0.804-0.630}{1-0.012}$

$= 0.176$

(b) Net Profit (in \$) $= \dfrac{X}{4} - 2$

$E(\text{Net Profit}) = \frac{1}{4}E(X) - 2 = \dfrac{20\times 0.2}{4} - 2 = -1$

$V(\text{Net Profit}) = \left(\frac{1}{4}\right)^2 V(X) = \dfrac{20\times(0.2)\times(0.8)}{4^2} = 0.2$

$SD_{\text{Net Profit}} = +\sqrt{0.2} = 0.4472$

(c) $E(\text{Net Profit}) > 0$

Let the instructor pay c cents.

$E(\text{Net Profit}) = E\left(\dfrac{cX}{100} - 2\right) = \dfrac{c}{100}\cdot E(X) - 2 = \dfrac{c\times(20)(0.2)}{100} - 2$

$= \dfrac{c}{25} - 2 > 0$

$\Rightarrow c > 50.$

Instructor should pay more than 50 cents.

("Fair game" if he pays exactly 50 cents)

(d) Let's call S' = An Exam with exactly 5 correct answers.

Y = # Exams required for the first S.

$p' = P(S') = P(X=5) = 0.174$

$Y \sim Geo(p')$

$P(Y=3) = (1-p)^2 p = (1-0.174)^2 (0.174)$

$= 0.119$

(e) $X \sim Bin\left(100, \frac{1}{25} = 0.04\right)$

$P(X \geq 2) = 1 - P(X \leq 1)$

$= 1 - \binom{100}{0}(0.04)^0(0.96)^{100} - \binom{100}{1}(0.04)^1(0.96)^{99}$

$= 0.913$

$n = 100 > 50, np = 4 < 5 \rightarrow$ Poisson approx. is valid

$X \overset{approx.}{\sim} Poi(\lambda = np = 4)$

$P(X \geq 2) = 1 - P(X \leq 1) \approx 1 - 0.092$ (Poisson Table: $x = 1, \lambda = 4$)

$= 0.908$

Approx. error $= 0.005$ ($= 0.55\ \%$)