Probability and Statistics

Session II

We begin with a question for homework

* Calculate $P(|X - \mu_X| \geq k\sigma_X)$ for

a) $X \sim$ uniform$(0,1)$

b) $X \sim$ exponential(λ)

You can try at least (a) before watching the rest of this lecture.

Common Families of Distributions

Discrete Distributions

(recall) $X \to$ Discrete dist. if Sample space of X is countable.

D. Uniform Dist.

X has a discrete uniform $(1, N)$ dist. if

$$P(X = x|N) = \frac{1}{N}, \quad x = 1, 2 \cdots, N$$

Integer

$\boxed{\text{Equal mass on all outcomes}}$

Notation '|' implies 'given'.

D. Uniform dist.

Mean

$$Ex = \sum_{x=1}^{N} x P(X=x \mid N)$$

$$= \sum_{x=1}^{N} x \frac{1}{N} = \frac{N+1}{2}$$

Variance

$$\mathrm{Var}X = Ex^2 - (Ex)^2$$

$$\vdots$$

$$= \frac{(N+1)(N-1)}{12}$$

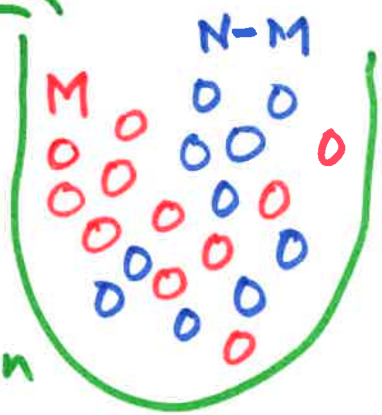
Identities

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2}$$

$$\sum_{i=1}^{k} i^2 = \frac{k(k+1)(2k+1)}{6}$$

D. Hypergeometric dist.

Take out K balls at random



Urn with N balls

then

what is the probability that n balls are Red.

$$* \quad P(X = n \mid N, M, K) = \frac{\binom{M}{n}\binom{N-M}{K-n}}{\binom{N}{K}}, \quad n = 0, 1 \ldots K$$

$$* \quad EX = \frac{KM}{N} \qquad * \quad \mathrm{Var}\,X = \frac{KM}{N}\left(\frac{(N-M)(N-K)}{N(N-1)}\right)$$

Proof — homework

D. Binomial dist.

X has Bernoulli (p) dist. if

$$X = \begin{cases} 1 & \text{with prob. } p \\ 0 & \text{with prob. } 1-p, \end{cases} \quad 0 \le p \le 1$$

Suppose n identical Bernoulli trials are done,

define $A_i = \{X = 1 \text{ on the } i\text{th trial}\}, i = 1, \cdots n.$

↑ collection of independent events

Y = total number of success in n trials

$$P(Y = y \mid n, p) = \binom{n}{y} p^y (1-p)^{n-y}, \qquad \text{Binomial}(n, p)$$

$y = 0, 1 \cdots n$

$\boxed{\text{Mean, Var} \rightarrow \text{Sec 2.4}}$

D. Poisson dist.

$X \rightarrow$ random variable taking values in non-negative integer

has a Poisson (λ) dist. if

$$P(X=x \mid \lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0, 1, \ldots$$

H.W. verify $\sum_{x=0}^{\infty} P(X=x \mid \lambda) = 1$ (hint: use Taylor Series)

$EX = \lambda$ $\quad$ $\text{Var}X = \lambda$ (easy to derive)

Poisson example and relation to Bernoulli trials

* A typesetter makes 1 error in every 500 words. A typical page is 300 words. What is the probability that there will be no more than 2 erros in 5 pages.

Soln: assume setting a word is a Bernoulli trial with $p = \frac{1}{500}$, then X = number of errors in 5 pages (1500 words) is binomial$\left(1500, \frac{1}{500}\right)$

$$P(X \leq 2) = \sum_{x=0}^{2} \binom{1500}{x} \left(\frac{1}{500}\right)^{x} \left(\frac{499}{500}\right)^{1500-x}$$

for Poisson dist.

$$P(X=x) = \frac{\lambda}{x} P(X=x-1) \quad \text{--- (1)}$$

for Binomial $\quad P(Y=y) = \frac{(n-y+1)}{y} \frac{p}{1-p} P(Y=y-1) \quad \text{--- (2)}$

if we set $\lambda = np$, $p \to$ small

$$\frac{(n-y+1)}{y} \frac{p}{1-p} \simeq \frac{\lambda}{y}$$

remain to show that

$P(X=0) \simeq P(Y=0)$ (try it out!)

hint : works for $\underline{\text{large } n}$.

$\longrightarrow$ go back to the previous example.

Other D. dist.

* Negative Binomial $P(X=x|r,p) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$

$x = r, r+1 \ldots$

* Geometric Distribution
(Special case of negative Binomial)

$$P(X=x|p) = p(1-p)^{x-1}, \quad x = 1, 2 \ldots$$

⋮

Homework:
a) Find the mean and variance of Geometric dist.
b) Think of an example in which Geometric dist can be used to model...

Continuous Distributions

C. Uniform dist.

$$f(x|a,b) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a,b] \\ 0 & \text{otherwise.} \end{cases}$$

$$EX = \frac{b+a}{2} \qquad \operatorname{Var} X = \int_a^b (\ ?\ )\, dx$$

HW

$$= \frac{(b-a)^2}{12}$$

Gamma Dist.

* Support - $[0, \infty)$

* Gamma function

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} \, dt$$

for α as integers

$$\Gamma(\alpha) = (\alpha-1)!$$

Think of gamma function as a generalization of the factorial function over real numbers

(Gamma dist.)

It can be verified that

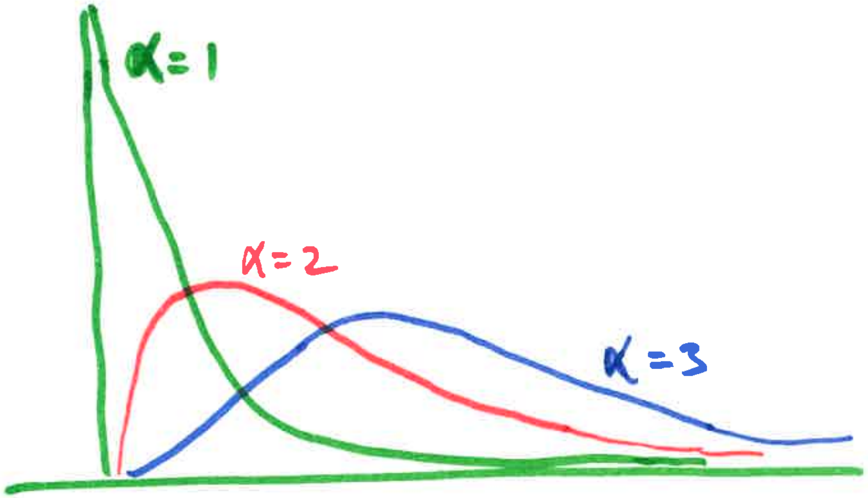
$$f(t) = \frac{t^{\alpha-1}e^{-t}}{\Gamma(\alpha)}, \quad 0 < t < \infty$$

$\downarrow$ change of variables $X = \beta T$

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, \quad 0 < x < \infty$$

$\alpha > 0$ (shape)

$\beta > 0$ (scale)



Mean and Variance of gamma dist.

$$EX = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x \underbrace{x^{\alpha-1} e^{-x/\beta}}_{\text{kernel of gamma}(\alpha+1,\beta)} dx$$

$$\vdots$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \Gamma(\alpha+1)\beta^{\alpha+1}$$

$$= \frac{\alpha \, \Gamma(\alpha)\beta}{\Gamma(\alpha)} = \alpha\beta$$

$\text{Var}\, X = \alpha\beta^2$ (can be verified similarly)

* Note on Gamma - Poisson Relationship when α is an integer

* Special Cases of gamma dist.

a) if $\alpha = p/2$, p is integer and $\beta = 2$

$$f(x|p) = \frac{1}{\Gamma(p/2)\, 2^{p/2}} x^{((p/2)-1)} e^{-x/2}, \quad 0 < x < \infty$$

<u>Chi squared pdf with 'p' degrees of freedom</u>

b) i) $\alpha = 1$ $\quad f(x|\beta) = \frac{1}{\beta} e^{-x/\beta}, \quad 0 < x < \infty$

<u>exponential pdf</u> with scale β

Normal distribution or Gaussian dist.

p.d.f.

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$$

* analytically tractable
* bell shape / symmetry – suitable for many population models
* Central Limit Theorem

(large variety of dist) + (large samples) $\xrightarrow[\text{conditions}]{\text{mild}}$ $\approx$ Normal distribution

* If $X \sim n(\mu, \sigma^2)$ then

$$Z = (X-\mu)/\sigma \sim \underbrace{n(0,1)}_{\text{Standard normal.}} \quad \text{— 1}$$

* $$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} \, dz = 1 \quad \text{— 2}$$

Homework : a) Show that 1 is true

b) Derive 2 but do not spend more than 1 hour on it.

Other continuous dist.

* Beta dist. $f(x|\alpha,\beta) = \dfrac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$,

$0 < x < 1$, $\alpha > 0$, $\beta > 0$

$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$

* Cauchy dist. $f(x|\theta) = \dfrac{1}{\pi}\dfrac{1}{1+(x-\theta)^2}$, $-\infty < x < \infty$
$-\infty < \theta < \infty$

* Lognormal $f(x|\mu,\sigma^2) = \dfrac{1}{\sqrt{2\pi}\sigma}\dfrac{1}{x}e^{-(\log x-\mu)^2/2\sigma^2}$,

$0 < x < \infty$
$-\infty < \mu < \infty$, $\sigma > 0$

* Double exponential
$f(x|\mu,\sigma) = \dfrac{1}{2\sigma}e^{-|x-\mu|/\sigma}$, $-\infty < x, \mu < \infty$
$\sigma > 0$

Multiple Random Variables

Univariate models vs.

↑ only 1 random var.

E.g. Samples of weights

$$S \longrightarrow \mathbb{R}$$

(Real num.)

Multivariate models

↑ many random variables

Eg: Samples of weight, height, B.P.

$$S \longrightarrow \mathbb{R}^n$$

n-dim Euc. Space.

n-dim Random Vector

Joint Distribution

(We focus on $\mathbb{R}^2$ first)

$(X\ Y) \longrightarrow$ discrete bivariate random vector

$$f_{X,Y}(x,y) : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$\underbrace{f_{X,Y}(x,y)}_{\text{joint PMF}} = P(X=x, Y=y)$$

Example: Toss of 2 fair dice

* 36 equally likely points
* let X = Sum of 2 dice

Y = |diff. of 2 dice|

$$\begin{array}{cc|ccccc} & & & x \longrightarrow & & & \\ & & 2 & 3 & \cdots & & 12 \\ \hline & 0 & 1/36 & \cdots & & \cdots & 1/36 \\ y\downarrow & 1 & & & & & \\ & \vdots & & & & & \\ & 5 & & & \frac{1}{18} & & \end{array}$$

Some Notes on joint PMF

* $$E\,g(X,Y) = \sum_{(x,y)\in\mathbb{R}^2} g(x,y)\,f(x,y)$$

* $$\sum_{(x,y)\in\mathbb{R}^2} f(x,y) = P\big((X,Y)\in\mathbb{R}^2\big) = 1$$

Marginal Distribution.

let $(X, Y) \rightarrow$ dis. bivariate r. v. and $f_{XY}(x,y)$ be joint PMF then

marginal PMF of X is

$f_X(x) = P(X = x)$ given by

$$f_X(x) = \sum_{y \in \mathbb{R}} f_{X,Y}(x,y)$$

(Same way for $f_Y(y)$)

It is the P~~B~~MF of X no matter what Y takes on as value, in the content of joint model

Continuous Case.

* $f_{X,Y}(x,y)$ from $\mathbb{R}^2$ to $\mathbb{R}$ is a joint PDF of the Continuous ~~r.v.~~ bivariate r.v. (X,Y) if for every $A \subset \mathbb{R}^2$

$$P\big((X,Y)\in A\big) = \iint_A f(x\,y)\,dx\,dy$$

* $$E\,g(X\,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\,f(x,y)\,dx\,dy$$

for all $(x,y)\in\mathbb{R}^2$

* Marginals $$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \quad , -\infty < x < \infty,$$

and Similarly for Y

Homeworks : (joint PDFs)

1) $$f_{X,Y}(x,y) = \begin{cases} 6xy^2 & 0<x<1 \text{ and } 0<y<1 \\ 0 & \text{otherwise} \end{cases}$$

Compute $P(X+Y \geq 1)$

2) $$f(x,y) = e^{-y} \, , \quad 0<x<y<\infty$$

Compute $P(X+Y \geq 1)$

Conditional Distribution.

$(X, Y) \rightarrow$ often x and Y can be related.

recall the weight & height exampl

$$f(y|x) = \frac{f(x,y)}{f_X(x)}$$

Conditional expected value of $g(Y)$

$$E(g(Y)|x) = \sum_{y} g(y)\, f(y|x)$$

$$E(g(Y)|x) = \int_{-\infty}^{\infty} g(y)\, f(y|x)\, dy$$

Independence

let $(X, Y) \rightarrow$ bivariate r.v. , PDF or PMF $f(x,y)$ and marginals $f_X(x)$ and $f_Y(y)$,

then X and Y are independent if for every $x \in \mathbb{R}$ and $y \in \mathbb{R}$

$$f(x,y) = f_X(x)\, f_Y(y) \;.$$

Also in case of independence

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = f_Y(y)$$

Bivariate Transformation.

let $(X, Y) \longrightarrow$ discrete bivarate. r.v.

define (U, V) as $U = g_1(X, Y)$ and $Y = g_2(X, Y)$

then

$$f_{U,V}(u,v) = P(U=u, V=v) = P\big((X,Y) \in A_{uv}\big) = \sum_{(x,y) \in A_{uv}} f_{XY}(x,y)$$

where

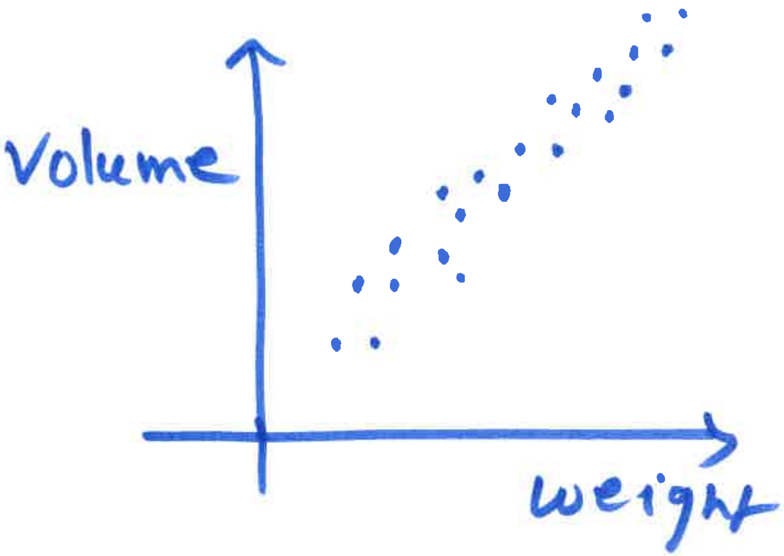
$$A_{uv} \stackrel{def}{=} \{(x,y) \in A : g_1(x,y) = u \ , \ g_2(x,y) = v\}$$

and A is a set where joint PMF of (X, Y) is positive.

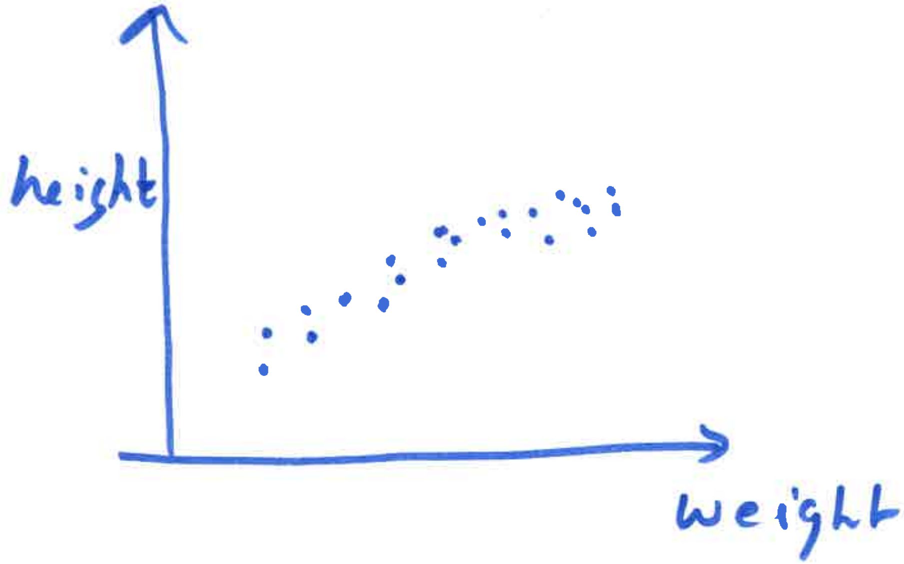
Covariance and Correlation

Is the relationship between dependent r.v.s Strong or weak?

Eg. 1 water samples



Eg. 2 human height-weight



Covariance

$$Cov(X,Y) = E\left((X-\mu_x)(Y-\mu_y)\right)$$

$$= EXY - \mu_x\mu_y$$

? Simple to try out yourself

Correlation

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_x \sigma_Y}$$

always between -1 and 1

Notes on Covariance

* Covariance of independent random variables is zero (H.W. show this formally)

* $$Var(aX + bY) = a^2 \, Var X + b^2 \, Var Y + 2ab \, Cov(X, Y)$$

(H.W. - derive this)

* recommended : Identities