WEEK 1

Random experiments → Outcome is random

Random experiments → Repeatable

Experiments

→ Raw Data (height of everyone)

→ Summarized raw data (avg. height)

How to do math for exp: probability model

Probability model

→ Sample space: set of all outcom.

→ Events: $\{A, B, C, \ldots\}$

→ Probability function: $P(A)$

↳ ① $P(A) \geq 0$

② $P(S) = 1$

③ If $A_1, \ldots, A_n$ mutually excl.

$$P(\cup A_i) = \sum P(A_i)$$

A and B and C (M.E)

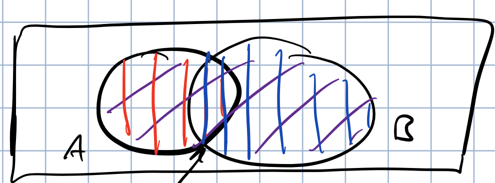
$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

- Prob. function properties:

① $P(\emptyset) = 0$

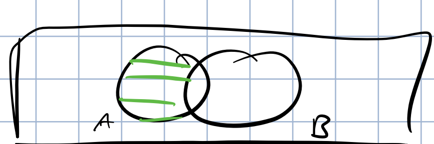
double counting

② $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



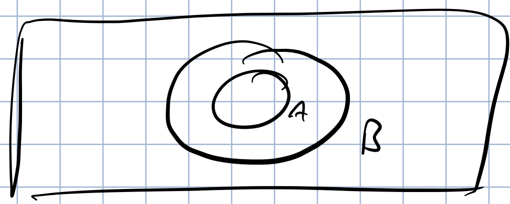
Double counting

③ $P(A \cap \overline{B}) = P(A) - P(A \cap B)$



④ $P(\bar{A}) = 1 - P(A)$

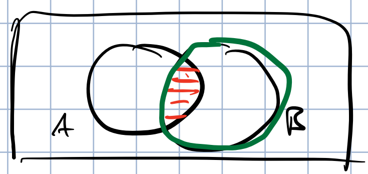
⑤ $A \subseteq B \Rightarrow P(A) \leq P(B)$



⑥ $0 \leq P(A) \leq 1$

- Conditional probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad (P(B) > 0)$$



- Independence: $P(A \cap B) = P(A)\,P(B) \Rightarrow A \perp B$

- Random variable:

X maps $S \to \mathbb{R}$

$0 \quad 1 \quad 3 \quad 2 \quad \ldots$

$\{A, B, C, D \ldots\}$

Valid r.v.: $\{X \leq x\} = \{A \in S : X(A) \leq x\}$

$X = \#$ heads in 2 coin flips:

$X = \{(T,T) \Rightarrow 0,\ (H,T) \wedge (T,H) \Rightarrow 1,\ (H,H) \Rightarrow 2\}$

$x \leq 1 \Rightarrow$

$\{X \leq 1\} = \{(H,T), (T,H)\}$

$\{\underbrace{A \in S : X(A) \leq x}_{\text{All domain values s.t. range} \leq 1 \text{ in } X}\} = \{(H,T), (T,H)\}$

R.V. → CDF

R.V. → PMF / PDF

CDF:

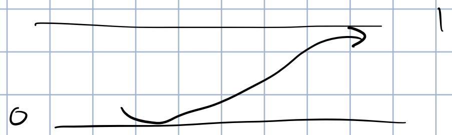
CDF of $X \Rightarrow F(x)$ like $F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}$

$$F(2) = P(X \leq 2)$$

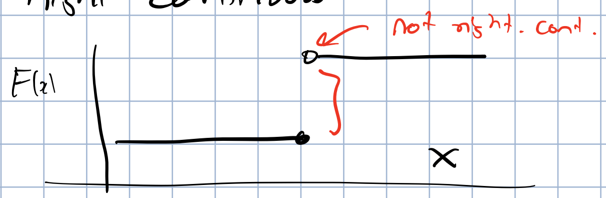
Properties:

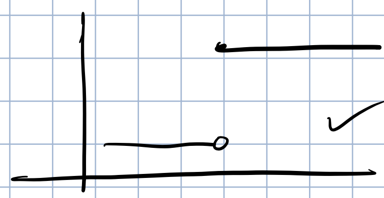
① F is non-decreasing

② $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$



③ Right continuous:



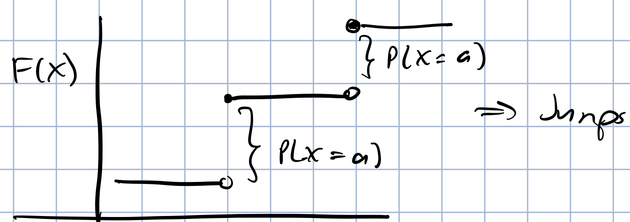


④ $P(a \le X \le b) = F(b) - F(a)$

⑤ $P(X = a) =$ Jump!

$$= \lim_{x \to a^+} F(x) - F(a)$$

Discrete r.v.: finite # of values.



PMF = $P(X = a) = f(a)$

Support: $A = \{x \in \mathbb{R},\ f(x) > 0\}$ → All values s.t. $P(a) > 0$

Properties:

① $f(x) \ge 0$

② $\sum_{x \in A} f(x) = 1$

→ Use these to check validity of pmf

PMF ⟷ CDF

PMF → CDF

$$F(x) = P(X \le x) = \sum_{t \in A,\, t \le x} f(t)$$

CDF → PDF

Use jump trick

Types of discrete r.v.

① Bernoulli

$X \sim \text{Bernoulli}(p) \Leftarrow X \in \{0, 1\} \wedge P(X=0) = 1-p, \; P(X=1) = p$

$$\text{pmf} = f(x) = \begin{cases} p^x (1-p)^x, & x \in (0,1) \\ 0 \end{cases}$$

② Binomial: repeated Bernoulli

$X \in \{0, 1\}$, independent + repeated trials, same prob. / trial

↳ $X \sim \text{Binomial}(n, p)$

B/c repeated Binomial

$$X = \sum_{i=1}^{n} X_i \quad (X_i \sim \text{Bernoulli}(p))$$

pmf: $$f(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

③ Geometric: binomial setup ⇒ # of times until 1^{st} success

$$f(x) = P(X=x) = (1-p)^x p$$

⇑

Probability of success w/in x failures.

④ Negative binomial: # of failures before r successes in repeated Bernoulli

$$f(x) = P(X=x) = \binom{x+r-1}{x} \underbrace{(1-p)^x}_{x \text{ failures}} \underbrace{p^{r-1}}_{r-1 \text{ success}} \underbrace{p}_{\text{next success}}$$

$$X = \sum_{i=1}^{r} X_i \quad (X_i \sim \text{Geo}(p))$$

⑤ Poisson: given rate μ, # of times event happen in the interval

↳ Poisson conditions:

1. Independent event
2. Proportional
3. Individual: events cannot happen at same.

$$f(x) = P(X=x) = \frac{\mu^x}{x!} e^{-\mu}$$

Continuous r.v.

How to know if continuous r.v.?

1. $F(x)$ is continuous $\forall x \in \mathbb{R}$
2. $F(x)$ is differentiable everywhere except for some finite points.

Prob. density function:

$$f(x) = F'(x) \quad (F(x) \text{ is diff.})$$

Support:

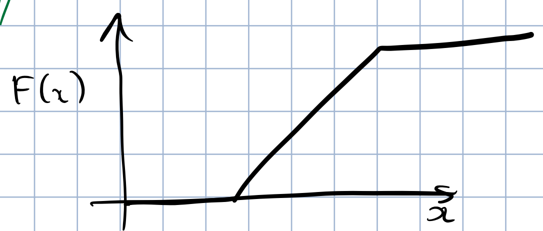
$$A = \{x \mid f(x) > 0\}$$

NOTE: $f(x) \neq P(X = x)$

Properties:

1. $f(x) \geq 0$
2. $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$
3. $F(x) = P(X \leq x) = \displaystyle\int_{-\infty}^{x} f(x)\,dx$
4. $P(a \leq X \leq b) = \displaystyle\int_{a}^{b} f(x)\,dx = F(b) - F(a)$
5. $P(X = a) = 0$

Ex://



$$F(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{x-b} & a \leq x < b \\ 1 & b < x \end{cases}$$

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0, & \text{o.w.} \end{cases}$$

WEEK 2

Gamma Function:

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\, dy \qquad \alpha > 0$$

Look @ converting CDF $\Rightarrow \Gamma(\alpha)$

Properties:

- $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$ $=$
- $\Gamma(n-1) = (n-1)!$ $\quad n = 1, 2, \ldots$
- $\Gamma(1/2) = \sqrt{\pi}$

Ex:// $X \sim N(0,1)$. $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$

Prove $\Sigma x = 1$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = 2\int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$$

$y = \frac{x^2}{2}$

$2y = x^2$

$x = \sqrt{2y}$

$dx = \frac{\sqrt{2}}{2} y^{-1/2} dy$

$$= \frac{2}{\sqrt{2\pi}} \int_0^\infty \exp(-y) \cdot \frac{\sqrt{2}}{2} y^{-1/2} dy$$

$$= \frac{1}{\sqrt{\pi}} \int_0^\infty y^{-1/2} \exp(-y)\, dy \quad \Rightarrow \quad \alpha - 1 = -1/2$$

$\alpha = -1/2 + 1$

$\alpha = 1/2$

$$= \frac{1}{\sqrt{\pi}} \Gamma(1/2)$$

$$= \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi}$$

$$= 1$$

Expectation

Discrete:

$$E[X] = \sum_{x \in A} x f(x) \quad \Rightarrow \text{Check if } \sum x f(x) \text{ converge}$$

Continuous:

$$\int_{-\infty}^{\infty} x f(x)\, dx$$

Ex://

$$X \Rightarrow f(x) = \begin{cases} \frac{1}{x(x+1)} , & x = 1, 2, \dots \\ 0 \end{cases}$$

$$\mathbb{E}(x) = \sum_{x \geq 1} x \cdot \frac{1}{x(x+1)} = \sum_{x \geq 1} \frac{1}{x+1}$$ $\Rightarrow$ Diverges, so $\mathbb{E}(x) = \text{DNE}$

Ex://

$$X \Rightarrow f(x) = \begin{cases} \frac{\theta}{x^{\theta+1}} & x \geq 1 \\ 0 & \text{o.w.} \end{cases}$$

1) Which values of θ does $E[x]$ exist

$$E[x] = \int_{-\infty}^{\infty} x \cdot \frac{\theta}{x^{\theta+1}} = \int_{-\infty}^{\infty} \frac{\theta}{x^{\theta}}$$ $\equiv$ Converges iff $\theta \geq 2$ $\hookrightarrow$ Harmonic sum.

2) Find $E[x]$

$$\mathbb{E}[x] = \theta \int_1^{\infty} \frac{1}{x^{\theta}} dx = \theta \cdot \frac{1}{-\theta+1} x^{-\theta+1} \Big|_1^{\infty}$$

$$= \theta \left(0 - \frac{1}{-\theta+1} \right)$$

$$= \theta \left(\frac{1}{\theta - 1} \right)$$

Take expectations of variables $(X^2, X+1, 2X \dots)$

$$\mathbb{E}[g(x)] = \sum g(x) f(x)$$

Linearity:

$$\mathbb{E}[a + bX] = a + b\,E[X]$$

Variance:

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

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

k^{th} moment:

$$E[X^k]$$

k^{th} central moment:

$$\mathbb{E}[(X - \mathbb{E}[X])^k]$$ $\Rightarrow$ Variance: 2^{nd} central moment.

Ex.// $X \sim \text{Uniform}(0,1)$

$$\mathbb{E}[X] = 1/2$$

$$\mathbb{E}[2X+1] = 2\mathbb{E}[X] + 1 = 2$$

$$\mathbb{E}[X^2] = \int_0^1 x^2 \cdot 1 \, dx = \frac{x^3}{3}\Big|_0^1 = 1/3$$

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$= 1/3 - (1/2)^2$$

$$= 1/3 - 1/4$$

Moment Generating Function

X r.v. $\Rightarrow$ $M(t) = \mathbb{E}[e^{tx}]$ $\forall t$ s.t. it exists

and $\exists h$ s.t. $\mathbb{E}[e^{tx}]$ exists in $t \in (-h, h)$ $\Rightarrow$ Construction of $M(t)$, not necessary

Ex.// $X \sim \text{Exp}(\theta)$

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & x > 0 \\ 0 & \text{o.w.} \end{cases}$$

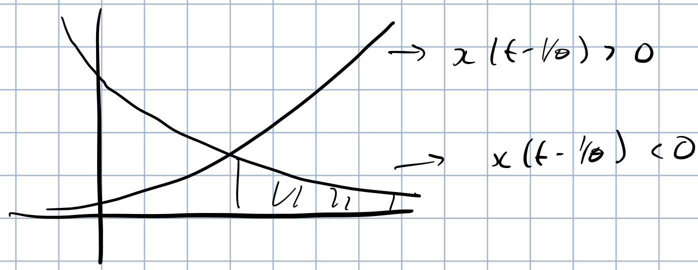
Find MGF

① When does $\mathbb{E}[e^{tx}]$ exist?

$$\mathbb{E}[e^{tx}] = \int_0^\infty e^{tx} \cdot \frac{1}{\theta} e^{-x/\theta} \, dx$$

$$= \frac{1}{\theta} \int_0^\infty e^{x(t - 1/\theta)} \, dx$$

As long as $x(t - 1/\theta) < 0 \Rightarrow t < 1/\theta$



② Find MGF

$$M(t) = \frac{1}{\theta} \int_0^\infty e^{x(t - 1/\theta)} \, dx \qquad \forall t < 1/\theta$$

$$= \frac{1}{\theta}\left[\frac{1}{t-1/\theta} e^{x(t-1/\theta)}\right]_0^\infty$$

$$= \frac{1}{1-t\theta}$$

Linearity of MGF:

X is a r.v. and domain $\mathcal{T}$. $Y = aX + b$.

$$M_Y(t) = e^{bt} M_X(at) \quad \forall t \in \{t \in \mathbb{R} : at \in \mathcal{T}\}$$

Properties:

① $M_X(0) = 1$

② $M_X^{(k)}(0) = \mathbb{E}[X^k]$ ⇒ k^{th} derivative of $M(0)$ gives k^{th} moment

Ex:// $X \sim Exp(\theta)$

$$M_X(t) = \frac{1}{1-t\theta} \quad \forall t < 1/\theta$$

Find the mean:

Mean is the 1^{st} moment: $\mathbb{E}[X] = \mathbb{E}[X^1]$

$$M_X(t)' = -\frac{1}{(1-t\theta)^2} \cdot (-\theta) \Big|_{t=0}$$

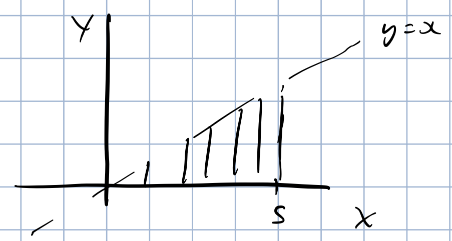
$$= \theta$$

MGF UNIQUELY DEFINES distributions (like CDF)

Joint and Marginal CDF

Joint CDF:

$$F(x, y) = P(X \leq x, Y \leq y) = P(\{X \leq x\} \cap \{Y \leq y\})$$



Marginal CDF:

$$F_X(x) = \lim_{y \to \infty} F(x, y) = F(X)$$

$$F_y(y) = \lim_{x \to \infty} F(x, y) = F(Y)$$

Properties.

1. F is non-decreasing if 1 variable fix

2. $\lim_{x \to -\infty} F(x,y) = \lim_{y \to -\infty} F(x,y) = 0$

3. $\lim_{(x,y) \to (-\infty,-\infty)} F(x,y) = 0$, $\lim_{(x,y) \to (\infty,\infty)} F(x,y) = 1$

Discrete Multivariate.

Countable # of pts.

Joint p.f.:

$$f(x,y) = P(X=x, Y=y)$$

Joint support:

$$A = \{(x,y) : f(x,y) > 0\}$$

Properties:

1. Non-negative $\forall x,y$

2. $\sum_{(x,y)\in A} f(x,y) = 1$

Marginal.

For x:

$$f(x) = \sum_{\text{all } y} f(x,y)$$

For y:

$$f(y) = \sum_{\text{all } x} f(x,y)$$

Ex.// $X, Y \Rightarrow$ r.v.

$$f(x,y) = \begin{cases} k(1-p)^2 p^{x+y} & x = 0,1,\ldots \quad y = 0,1,2,\ldots \\ 0 & \end{cases}$$

a) Find k:

Relate to properties.

① Positive: $k \geq 0$

② Sum to 1:

$$\sum_{x=0}^{\infty} \sum_{y=0}^{\infty} k(1-p)^2 p^{x+y} = 1$$

$$k(1-p)^2 \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} p^{x+y} = 1$$

$$k(1-p)^2 \sum_{x=0}^{\infty} p^x \sum_{y=0}^{\infty} p^y = 1$$

$$k\cancel{(1-p)^2} \cdot \frac{1}{\cancel{1-p}} \cdot \frac{1}{\cancel{1-p}} = 1$$

$$k = 1$$

1. Separated
   ↳ Constants
   ↳ $x, y \Rightarrow$ resp. sums
2. Geometric sums

b) Marginal pdf

$$f_X(x) = \sum_{y=0}^{\infty} (1-p)^2 p^{x+y} = (1-p)^2 p^x \sum_{y=0}^{\infty} p^y$$

$$= (1-p)^2 p^x \cdot \frac{1}{1-p}$$

$$= (1-p)p^x$$

c) $P(X \leq Y)$

$$\sum_{x=0}^{\infty}\sum_{y=x}^{\infty} (1-p)^2 p^{x+y}$$

$$= (1-p)^2 \sum_{x=0}^{\infty} p^x \sum_{y=x}^{\infty} p^y$$

$$= (1-p)^2 \cdot \sum_{x=0}^{\infty} \cdot (p^x + p^{x+1} + p^{x+2} + \dots)$$

$$= (1-p)^2 \cdot \sum_{x=0}^{\infty} p^x (p^x \cdot (1 + p^1 + p^2 + \dots))$$

Useful trick when starting off at index $\neq 0$

$$= (1-p) \cdot \sum_{x=0}^{\infty} p^x \cdot p^x \sum_{y=0}^{\infty} p^y$$

$$= (1-p)^2 \cdot \sum_{x=0}^{\infty} p^{2x} \sum_{y=0}^{\infty} p^y$$

$$= (1-p)^2 \cdot \frac{1}{1-p^2} \cdot \frac{1}{(1-p)}$$

$$= \cancel{(1-p)^2} \cdot \frac{1}{(1+p)\cancel{(1-p)}} \cdot \frac{1}{\cancel{1-p}}$$

$$= \frac{1}{1+p}$$

Continuous Multivariate

$F(x,y)$ is continuous $\wedge$ $\frac{\partial^2}{\partial x \partial y} F(x,y)$ (except at countable pts of discontinuity) $\Rightarrow$ continuous.

Support: $A = \{(x,y) : f(x,y) > 0\}$

Properties:

① $f(x,y) \geq 0$

② $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

③ $P((X,Y) \in R)$

$$\iint_R f(x,y)\,dx\,dy$$

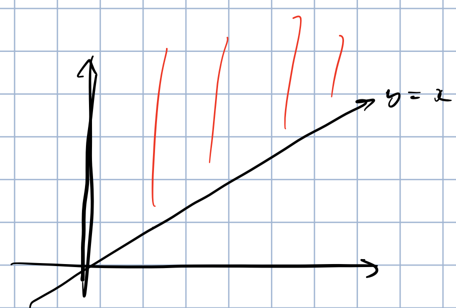
↳ $P(X \leq Y)$: $$\int_{-\infty}^{\infty}\int_{x}^{\infty} f(x,y)\,dy\,dx$$

Joint pdf → marginal: integrate on other variable

Ex://

$$f(x,y) = \begin{cases} ke^{-x-y} & 0<x<y<\infty \\ 0 & o.w. \end{cases}$$

① Draw area



a) Find k:

$$\iint_{Region} f(x,y)\,dy\,dx = 1$$

$$k\int_0^{\infty}\int_x^{\infty} e^{-x-y}\,dy\,dx = k\int_0^{\infty} -e^{-x-y}\Big|_x^{\infty}\,dx$$

$$= k\int_0^{\infty} 0-(-e^{-2x})\,dx$$

$$= k\int_0^{\infty} e^{-2x}\,dx$$

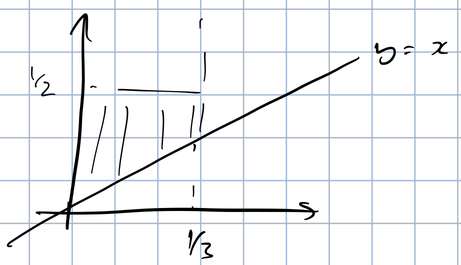
$$= k\left[-\tfrac{1}{2}e^{-2x}\right]_0^{\infty}$$

$$= k\left(0-\left(-\tfrac{1}{2}\times 1\right)\right)$$

$$1 = k/2$$

$$\boxed{k = 2}$$

b) $P(X \leq 1/3,\ Y \leq 1/2)$



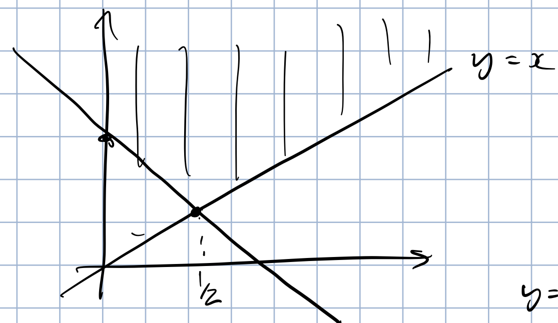
$$P(X \leq 1/3,\ Y \leq 1/2) = \int_0^{1/3}\int_x^{1/2} 2e^{-x-y}\,dy\,dx$$

c) $P(X \geq Y) = 1$ by defn.

d) $P(X + Y > 1)$

$X + Y > 1$

$Y > -X + 1$ $\Rightarrow$



$y = x$

$-x + 1 = x$

$2x = 1$

$x = 1/2$

$$P(X + Y > 1) = 1 - P(X + Y \leq 1)$$

$$= 1 - \int_0^{1/2}\int_x^{1-x} 2e^{-x-y}\,dy\,dx$$

e) Find marginal CDF of Y

$$F_Y(y) = \int_0^{\infty} 2e^{-x-y}\,\underline{\underline{dx}}$$

Independent R.V.

Event indep.:

$$P(A \cap B) = P(A)P(A) \Leftrightarrow A, B \text{ are indep.}$$

R.V.: A, B are indep. iff $\forall A, B \subseteq \mathbb{R}$

$$P(X \in A,\ Y \in B) = P(X \in A)P(Y \in B)$$

Indep. Theorem: X, Y are r.v. Prove indep. in 2 ways

① $F_1(x)$, $F_2(y)$: X, Y indep. $\Leftrightarrow F(x,y) = F_1(x)\,F_2(y)$

② $f_1(x), f_2(y)$: X, Y indep. $\Longrightarrow$ $f(x,y) = f_1(x) f_2(y)$

$\forall (x,y) \in A \times B$

Q: What about supports??

If $F(x,y)$ or $f(x,y)$ def. on support $A \times B$, f_1 defn on A, f_2 def. on B.

Ex://

$$f(x,y) = \begin{cases} (1-p)^2 p^{x+y} & x = 0, 1, \ldots \\ & y = 0, 1, \ldots \\ 0 & \text{o.w.} \end{cases}$$

Is X and Y indep.?

① Find marginal

$f_1(x) = (1-p)p^x$ ← $A = \{0, 1, \ldots\}$

$f_2(y) = (1-p)p^y$ ↳ $B = \{0, 1, \ldots\}$

$\Rightarrow f(x,y)!$ $\Rightarrow$ Def. on $A \times B$

② Prove

$\therefore$ Indep.

Q: Do you need to find marginals? NO!

Extension to indp. theorm

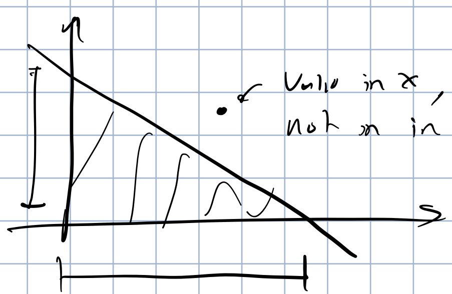
Factorize theorm

X, Y are indp. $\Longleftrightarrow$ $\exists\, g(x), h(y)$ st. $f(x,y) = g(x)h(y)$ ← Two random function!

$\forall (x,y) \in A_1 \times A_2$

$g(x) \propto f_1(x)$, $h(y) \propto f_2(y)$

Extension: if joint support not rect., not indp.



Valid in X, Valid in Y, not in $f(X, Y)$!

Ex://

$$f(x,y) = \begin{cases} \dfrac{\theta^{x+y} e^{-2\theta}}{x!\, y!} & x = 0, 1, \ldots \\ & y = 0, 1, \ldots \\ 0 & \text{o.w.} \end{cases}$$

→ Rectangular co-support

Is X, Y indp?

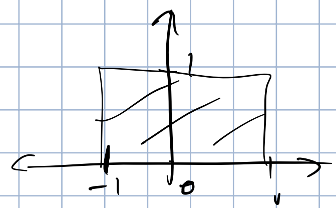
$$f(x,y) = \frac{\theta^{x+y} e^{-2\theta}}{x!\,y!} = \underbrace{\frac{e^{x} e^{-\theta}}{x!}}_{g(x)} \cdot \underbrace{\frac{e^{y} e^{-\theta}}{y!}}_{h(y)}$$

* Any conditions to check? $g(x), h(y) \geq 0 \quad \forall x, y$ *

Ex://

$$f(x,y) = \begin{cases} \frac{3}{2} y \, (1-x^2) &, -1 < x \leq 1, \quad 0 < y < 1 \\ 0 &, \text{o.w.} \end{cases}$$

① Joint support rect.?



② Factorize;

$$f(x,y) = \underbrace{\left(\tfrac{3}{2} y\right)}_{g(x) \geq 0} \underbrace{\left(1-x^2\right)}_{h(y) \geq 0}$$

$$-1 \leq x \leq 1 \qquad\qquad 0 < y < 1$$

Find marginal?

$$f_1(x) = \int_{-1}^{1} k(1-x^2)\, dx = 1$$

$$f_2(y) = \int_{0}^{1} k\left(\tfrac{3}{2} y\right) dy = 1$$

* NOTE: X, Y are indep. $\Rightarrow$ $g(X), h(Y)$ are also indep.!!

WEEK 4

Joint Expectations

① Define expectations

X

↳ Discrete: $E[X] = \sum^{\infty} x f(x)$

↳ Continuous: $\int_A x f(x)\, dx$

→ 1 variable

Multivariate expectations.

Ⓐ Discrete

$$\sum_{x=0}^{\infty} \sum_{y=0}^{\infty} |h(x,y)| f(x,y) < \infty \Rightarrow E[h(x,y)] = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} |h(x,y)| f(x,y)$$

pmf

Func. that we take expe.

Ⓑ Continuous:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |h(x,y)| f(x,y) < \infty \Rightarrow E[h(x,y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) f(x,y)\, dx\, dy$$

Ex:// $E[2X+3Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (2x+3y) f(x,y)\, dy\, dx$

Properties:

i) Linearity:

$$E[a h(x,y) + b g(x,y)] = a E[h(x,y)] + b E[g(x,y)]$$

ii) Indep of expectations.

$$X, Y \perp \Rightarrow \forall\, g(\cdot), h(\cdot) \quad E[g(x) h(y)] = E[g(x)] E[h(y)]$$

② Covariance: correlation of 2 variable

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

$$Cov(X, Y)$$

$<$ Opposite corr. | $=$ No corr. | $>$ Correlated

Properties:

1) Convenience: $Cov(X, Y) = E[XY] - E[X]E[Y]$

2) Independence: $Cov(X, Y) = 0$

3) Linearity of variance:

Single var: $Var(X) = E[X^2] - (E[X])^2$

Multiple variables:

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

Extend:

$$Var\left(\sum_{i=0}^{n} a_i x_i\right) = \sum_{i=0}^{n} a_i^2 Var(x_i) + \underbrace{\sum_{i \neq j} a_i a_j Cov(x_i, x_j)}_{0 \text{ if } x_i \perp x_j}$$

4) Weird property:

$$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$$

Ex: // X, Y are r.v.

$$f(x, y) = \begin{cases} x + y, & 0 < x < 1,\ 0 < y < 1 \\ 0, & \text{o.w.} \end{cases}$$

Find $Var[X + Y]$.

① Break it down via linearity

$$Var[X + Y] = Var(X) + Var(Y) + 2\, Cov(X, Y)$$

② Variance of X $\Rightarrow$ expectations $\Rightarrow$ marginal

i) Marginal of X:

$$f_X(x) = \int_0^1 (x + y)\, dy = xy + \frac{y^2}{2}\Big|_0^1 = x + \tfrac{1}{2}$$

ii) Expectation:

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

$$E[x^2] = \int_0^1 x^2 (x+1/2)\,dx = \frac{x^4}{4} + \frac{x^3}{6} \Big|_0^1 = \frac{5}{12}$$

$$E[x] = \int_0^1 x(x+1/2)\,dx = \frac{x^3}{3} + \frac{x^2}{4} \Big| = \frac{7}{12}$$

iv) Variance:

$$Var(x) = E(x^2) - (E(x))^2$$
$$= \frac{5}{12} - \left(\frac{7}{12}\right)^2$$
$$= \frac{11}{144}$$

③ Variance of Y:

$$Var(Y) = \frac{11}{144}$$

④ Cov (X, Y):

$$Cov(X,Y) = E[XY] - E[X]\,E(Y)$$

$$E(XY) = \int_0^1 \int_0^1 xy\,(x+y)\,dy\,dx = 1/3$$

$$Cov(X,Y) = \frac{1}{3} - \left(\frac{7}{12}\right)^2$$
$$= -\frac{1}{144}$$

⑤ Calculate:

$$Var(X+Y) = \frac{11}{144} \times 2 - \frac{1}{72}$$
$$= \boxed{\frac{5}{72}}$$

③ Correlation coefficient

Covariance is <u>not</u> normalized $\Rightarrow$ hard to compare.

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)}\,\sqrt{Var(Y)}} \Rightarrow -1 \le \rho \le 1$$

Ex:// Corr. coef. of prev. example.

$$\rho = \frac{-1/144}{\sqrt{\frac{11}{72}}\sqrt{\frac{11}{72}}} = \boxed{-1/11}$$

Conditional Distribution

① Definition:

Conditional pmf/pdf of X given $Y=y$ is:

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

Joint dist.

Univariate:

$$f(A|B) = \frac{P(A\cap B)}{P(B)}$$

Conditional pmf/pdf of Y given $X=x$:

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

Marginal,

② Properties:

1) $f_1(x|y) \geq 0 \quad \forall x, y$

2) $\sum f_1(x|y) = 1$

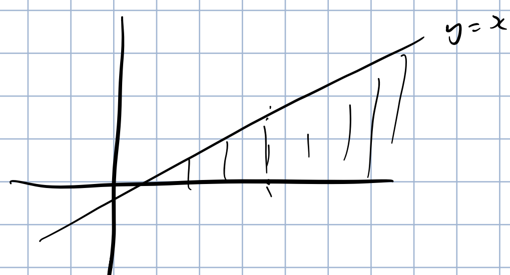
3) Calculation:

$$P(Y \leq y | X=x) = \int_{-\infty}^{y} f_2(t|x)\,dt$$

$$P(X \leq x | Y=y) = \int_{-\infty}^{x} f_1(t|y)\,dt$$

Ex://

$$f(x,y) = \begin{cases} 8xy, & 0<y<x<1 \\ 0, & \text{o.w.} \end{cases}$$



Calculate conditionals.

① Find marginal dist.

$$f_1(x) = \int_0^x 8xy\,dy$$

$$= \left[8x\frac{y^2}{2}\right]_0^x$$

$$= 4x^3$$

$$f_2(y) = \int_y^1 8xy\,dx = 4y - 4y^3$$

② Conditional calculation.

$$f_1(x|y) = \frac{f(x,y)}{f_2(y)} = \frac{8xy}{4y-4y^3} = \frac{2x}{1-y^2} \Rightarrow 0<y<x<1$$

$$f_2(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{8xy}{4x^3} = \frac{2y}{x^2}$$

4) Independence:

$$X \perp Y \Leftrightarrow f_1(x|y) = f_1(x),\quad f_2(y|x) = f_2(y),\quad \forall x,y.\ x \in A_1,\ y \in A_2$$

③ Product rule: use 1 marginal + 1 conditional $\Rightarrow$ joint distr.

$$f(x,y) = f_1(x|y)\cdot f_2(y) = f_2(y|x)\cdot f_1(x)$$

Note: Exactly using independence rule

$$f(x|y) = \frac{f(x,y)}{f_2(y)} \Rightarrow f(x|y)\,f_2(y) = f(x,y)$$

Really useful $\Rightarrow$ find other marginals/conditionals using joint distr.

Ex:// $Y \sim Poi(\lambda)$, $X|Y=y \Rightarrow$ Binomial (y,p). Distribution of X?

Need to find marginal of X.

① Denote variables:

$y = 0, 1, \ldots$

$$Y = \frac{\lambda^y e^{-\lambda}}{y!} \qquad X|_{Y=y} = \binom{y}{x} p^x (1-p)^{y-x}$$

$y = 0, 1, 2 \ldots$
$x = 0, 1, \ldots, y$

② Find joint distr. via product rule

$$F(X|_{Y=y})\,f(y) = \binom{y}{x} p^x (1-p)^{y-x} \cdot \frac{\lambda^y e^{-\lambda}}{y!}$$

$$= \frac{\cancel{y!}}{x!(y-x)!} p^x (1-p)^{y-x} \cdot \frac{\lambda^y e^{-\lambda}}{\cancel{y!}}$$

$$= \frac{\lambda^y p^x e^{-\lambda} (1-p)^{y-x}}{x!(y-x)!}$$

③ Find marginal

$$\sum_{y=0}^{\infty} f(x,y) = \sum_{y=x}^{\infty} \frac{\lambda^y p^x e^{-\lambda} (1-p)^{y-x}}{x!\,(y-x)!}$$

$$= \frac{p^x e^{-\lambda}}{x!} \sum_{y=x}^{\infty} \frac{\lambda^y (1-p)^{y-x}}{(y-x)!}$$ ⇒ Poisson: $\frac{\lambda^x e^{x}}{x!}$

$$= \frac{p^x e^{-\lambda}}{x!} \sum_{u=0}^{\infty} \frac{\lambda^{u+x} (1-p)^{u}}{u!}$$ $u = y - x$

$$= \frac{p^x e^{-\lambda} \lambda^x}{x!} \sum_{u=0}^{\infty} \frac{\lambda^{u} (1-p)^{u}}{u!}$$ ⇒ Exponential: $\sum \frac{y^x}{x!} = e^y$

$$= \frac{p^x e^{-\lambda} \lambda^x}{x!} \cdot e^{\lambda(1-p)}$$

$$= \frac{(p\lambda)^x e^{-\lambda p}}{x!} \qquad x = 0, 1, \ldots$$

$\therefore \quad X \sim Poi(p\lambda)$

WEEK 5

Conditional Expectation

$$E[g(Y) \mid X = x] = \begin{cases} \sum_{\text{all } y} g(y) \boxed{f_2(y \mid x)} \\ \int_{-\infty}^{\infty} g(y) f_2(y \mid x)\, dy \end{cases}$$

→ x is considered a constant! Only sum over y / variable

Special case:

1. $g(Y) = Y$

↳ $E[Y \mid X = x]$ ⇒ Conditional mean

2. $g(Y) = (Y - E[Y \mid X = x])^2$ ⇒ Conditional variance

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

↳ $Var[Y \mid X = x] = E[(Y - E[Y \mid X = x])^2 \mid X = x]$

$$= \underline{E[Y^2 \mid X = x] - (E[Y \mid X = x])^2}$$

Ex://

$$f(x,y) = \begin{cases} 8xy & 0<y<x<1 \\ 0 & \text{o.w.} \end{cases}$$

$$f_1(x|y) = \frac{8x}{4-4y^2}, \quad \underline{0<y<x<1}, \qquad f_2(y|x) = \frac{2y}{x^2} \quad 0<y<x<1$$

$E[X|Y=y]$? $\quad$ $Var[X|Y=y]$?

a) $E[X|Y=y]$

Apply expectation formula

$$E[X|Y=y] = \int_y^1 x \cdot f_1(x|y) = \dots = \frac{2}{3(1-y^2)} - \frac{2y^3}{3(1-y^2)}$$

b) $Var[X|Y=y]$?

$$Var[X|Y=y] = E[X^2|Y=y] - (E[X|Y=y])^2$$

$$E[X^2|Y=y] = \int_y^1 x^2 f_1(x|y) = \frac{1}{2}(1+y^2)$$

Define support of expectation via support of variable

$$Var[X|Y=y] = \frac{1}{2}(1+y^2) - \frac{4}{9}\left(\frac{1-y^3}{1-y^2}\right)^2 \qquad \boxed{0<y<1}$$

Special rules:

① Independence:

Condition has no effect

$$X \perp Y \Rightarrow E[g(X)|Y=y] = E[g(X)]$$

$$Var[g(X)|Y=y] = Var[g(X)]$$

② Substitution rule $\quad$ b/c $X = x$ from conditional distr.

$$E[h(X,Y)|X=x] = E[h(x,Y)|X=x] \qquad \boxed{\text{r.v} \Rightarrow \text{constant}}$$

Ex:// $E[X+Y|X=x] = E[x+Y|X=x]$ => LINEARITY

$= E[x|X=x] + E[Y|X=x]$

$= x + E[Y|X=x]$

Ex:// $E[XY|X=x] = E[xY|X=x] = x\,E[Y|X=x]$

③ All other properties of expectation holds

④ Double expectation:

Distinction:

$$E[g(X)|Y] \neq E[g(X)|Y=y]$$

Y is not fixed!

Constant

Y is fixed

$\therefore E[g(X)|Y] = h(Y)$

Formula:

Figure out expectation of 1 var from conditional!!

$$E[E[g(x)|Y]] = E[g(x)]$$

Special case: $g(X) = X \Rightarrow$ Law of Total Exp.

Ex:// $Y \sim Poi(\lambda)$ $\quad X|Y=y \sim Bin(y,p)$. $\quad X \sim Poi(p\lambda)$

$E[X|Y=y] = yp$

$E[X|Y] = Yp$

$E[E[X|Y]] = E[Yp] = p\,E[Y]$

$= p\lambda$

$E[X] = p\lambda$

⑤ Variance formula using conditional expe.

Old formula: $Var[Y] = E[Y^2] - (E[Y])^2$

What happens if you don't know Y and only know conditional

$$Var[Y] = E[Var[Y|X]] + Var[E[Y|X]]$$

Calculation steps:

1. $E[Var[Y|X]]$

   i) Find $Var[Y|X=x]$ $\Rightarrow$ substitute $X=x$

   ii) Take $E[h(x)]$

2. $Var[E[Y|X]]$

   i) Find $E[Y|X] \Rightarrow h(X)$

   ii) Calculate variance.

Ex:// $Y|X=x \sim Bin(x,p)$. $X \sim Poi(\lambda)$

What's the variance of Y. What's $E[Y]$

① $E[Var[Y|X]]$

i) $h(x) = Var[Y|X=x]$

$= xp(1-p)$

$h(X) = Xp(1-p)$

ii) $E[Var[Y|X]] = E[Xp(1-p)]$

$= p(1-p)E[X]$

$= p(1-p)\lambda$

$E[Y] = E[E[Y|X]]$

$= E[Xp]$

$= pE[X]$

$= p\lambda$

② $Var[E[Y|X]]$

i) $h_2(x) = E[Y|X=x]$

$= xp$

$h_2(X) = Xp$

ii) $Var[E[Y|X]] = Var[Xp] = p^2 Var[X]$

$= p^2\lambda$

③ $Var[Y] = p\lambda - p^2\lambda + p^2\lambda$

$= p\lambda$

Joint MGFs

$E[e^{t_1 X + t_2 Y}]$ exists $\forall\, t_1 \in (-h_1, h_1),\ t_2 \in (-h_2, h_2) \Rightarrow M(t_1, t_2) = E[e^{t_1 X + t_2 Y}]$

↳ Extend:

$$M(t_1, \ldots, t_n) = E\left[e^{\sum_{i=1}^{n} t_i x_i}\right]$$

Ex://

$$f(x,y) = \begin{cases} e^{-x-y} & x>0,\ y>0 \\ 0 & \text{o.w.} \end{cases}$$

Find joint MGF

① Existance:

$$E[e^{t_1 X + t_2 Y}] = \int_0^\infty \int_0^\infty e^{t_1 x + t_2 y} f(x,y)\, dy\, dx$$

$$= \int_0^\infty \int_0^\infty e^{t_1 x + t_2 y} e^{-x-y}\, dy\, dx$$

Find 2 cond. for both t_1, t_2. Break up integral to analyze sep. $\Rightarrow$

$$= \int_0^\infty \left(\int_0^\infty (e^{t_1 x - x})\, dx \right) e^{t_2 y - y}\, dy$$

Converges if $t_1 - 1 < 0$, $t_1 < 1$

↳ Converges if $t_2 - 1 < 0$, $t_2 < 1$

$\therefore E[e^{t_1 X + t_2 Y}]$ exists if $t_1 \in (-1/2, 1/2)$, $t_2 \in (-1/2, 1/2)$

② Calculate

$$M(t_1, t_2) = E[e^{t_1 X + t_2 Y}] = \int_0^\infty e^{t_2 y - y} \left(\int_0^\infty e^{t_1 x - x}\, dx \right) dy$$

$$= \left(\int_0^\infty e^{t_2 y - y}\, dy \right) \left(\int_0^\infty e^{t_1 x - x}\, dx \right)$$

$$(e^{t_2 y - y})' = (t_2 - 1) e^{(t_2 y - y)}$$

$$= \frac{e^{y(t_2 - 1)}}{t_2 - 1} \Bigg|_0^\infty \frac{e^{x(t_1 - 1)}}{t_1 - 1} \Bigg|_0^\infty$$

$$= \left(\frac{1}{1 - t_2} \right) \left(\frac{1}{1 - t_1} \right)$$

Can we find marginal MGF? Yes!

$$M(t_1) = M(t_1, \cancel{t_2}^{0}) = E[e^{t_1 x}]$$

Uniqueness of MGF: MGF uniquely defines distr!

Independence:

$$X \perp Y \iff M(t_1, t_2) = M_x(t_1) M_y(t_1)$$

Ex:// $X \sim Poi(\lambda_1)$, $Y \sim Poi(\lambda_2)$ and $X \perp Y$. Show $X+Y \sim Poi(\lambda_1 + \lambda_2)$

Idea: if we can show MGF of $(X, Y) \Rightarrow$ find distribution.

① Joint MGF:

$$M(t_1, t_2) = M_x(t_1) M_y(t_2)$$
$$= e^{\lambda_1(e^{t_1}-1)} e^{\lambda_2(e^{t_2}-1)}$$
$$= e^{\lambda_1(e^{t_1}-1) + \lambda_2(e^{t_2}-1)}$$

② Show $M(t_1, t_2) \Rightarrow Poi(\lambda_1 + \lambda_2)$

$$M(t, t) = E[e^{t(X+Y)}] \Rightarrow \text{Manipulated MGF into } X+Y$$
$$= e^{\lambda_1(e^t - 1) + \lambda_2(e^t - 1)}$$
$$= e^{\overbrace{(e^t - 1)}^{\text{MGF var.}} \overbrace{(\lambda_1 + \lambda_2)}^{\lambda}}$$

↳ MGF of Poisson (λ_1, λ_2)!

$$A \sim Poi(\lambda) \Rightarrow M(t) = e^{(e^t - 1)\lambda}$$
$$\boxed{Poi(\lambda_1 + \lambda_2)} \Rightarrow M(t) = \underline{e^{(e^t - 1)(\lambda_1 + \lambda_2)}}$$

We want MGF of $X+Y$ to look like that!

MGF of $X+Y$

$M(t) = E[e^{t(X+Y)}]$!!

WEEK 6

Multivariate distributions

Multinomial

Bivariate Normal

Multinomial

Extended version of binomial

↳ $X \sim \text{Multinomial}(n, p_1, \ldots, p_k) \Rightarrow$ # of times outcome i is picked out of k outcomes.

Ex:// K boxes, choose boxes randomly at p_i. # of times I choose i^{th} box in n trials

$$(X_1, X_2, \ldots, X_k) \sim \text{Multinomial}(n, p_1 \ldots, p_k)$$

$$f(x_1, \ldots, x_k) = \begin{cases} \dfrac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_2^{x_k} \\ 0, \quad \text{o.w.} \end{cases}$$

Condition:

① $x_i = 0, 1, \ldots$

② $\sum_{i=1}^{k} x_i = n$

Special case:

1. $k = 2 \Rightarrow$ Binomial!

Properties:

1. MGF.

$$M(t_1, \ldots t_k) = (p_1 e^{t_1} + \cdots + p_k e^{t_k})$$

2. Marginally: $X_i \sim \text{Binomial}(n, p_i)$

Pf Proof:

$$M_{X_i}(t_i) = M(0, 0, \ldots, t_i, 0 \ldots)$$

$$= p_i e^{t_i} + \sum_{i \neq j} p_j$$

$$= p_i e^{t_i} + 1 - p_i \Rightarrow \text{MGF of Binomial}$$

3. $T = X_i + X_j \sim \text{Bin}(n, p_i + p_j)$

4. $E[X_i] = np_i$, $Var[X_i] = np_i(1-p_i)$

$Cov(X_i, X_j) = -np_ip_j$

5. $X_i \mid X_j = x_j \sim Bin\left(n - x_j, \frac{p_i}{1-p_j}\right)$

6. $X_i \mid X_i + X_j = t \sim Bin\left(t, \frac{p_i}{p_i + p_j}\right)$

Bivariate Normal Distribution:

$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim BVN(\vec{\mu}, \Sigma)$

- $\vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \in \mathbb{R}^2$
- $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \Rightarrow$ Covariance matrix!

  $\hookrightarrow \sigma_1, \sigma_2 > 0$, $-1 < \rho < 1$

PDF:

$$f(\vec{x}) = \frac{1}{2\pi(|\Sigma|)^{1/2}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1}(\vec{x}-\vec{\mu})\right)$$

det

Properties:

1. MGF: $M(t_1, t_2) = e^{\vec{t}^T\vec{\mu} + \frac{1}{2}\vec{t}^T\Sigma\vec{t}}$

2. Marginally normal!

$$X_1 \sim N(\mu_1, \sigma_1^2) \quad , \quad X_2 \sim N(\mu_2, \sigma_2^2)$$

3. Conditional:

$$X_2 \mid X_1 = x_1 \sim N\left(\mu_2 + \rho\sigma_2\sigma_1^{-1}(x_1 - \mu_1), \sigma_2^2(1-\rho^2)\right)$$

4. Linearity:

$$\vec{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \neq \vec{0} \Rightarrow \vec{c}^T \vec{X} \sim N(\underbrace{\vec{c}^T \vec{\mu}}_{\text{Mean multiply}}, \underbrace{\vec{c}^T \Sigma \vec{c}}_{\text{Covariance multiply}})$$

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \ |A| \neq 0, \ \forall b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \Rightarrow A\vec{X} + \vec{b} \sim BVN(A\vec{\mu} + \vec{b}, A \Sigma A^T)$$

5. Facts:

$E[X_1] = \mu$. $Var[X_1] = \sigma_1^2$

$Cov(X_1, X_2) = \rho \sigma_1 \sigma_2$ $\quad Corr(X_1, X_2) = \rho$

6. Unique property of BVN

$$Cov(X_1, X_2) = 0 \Leftrightarrow \rho = 0 \Leftrightarrow X_1 \perp X_2$$

7. Independence:

$$X_1 \sim N(\mu_1, \sigma_1^2), \ X_2 \sim N(\mu_2, \sigma_2^2) \wedge X_1 \perp X_2$$

$$\therefore \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim BVN\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right)$$

TRANSFORMATIONS OF R.V.

Q: $Y = h(X_1, \ldots, X_n) \Rightarrow$ distribution

- $Y = X_1 X_2^2$, $Y = \sqrt{X_1}$

Methods

CDF:
CDF of Y | CDF of $X_1 \ldots X_n$

1-1 transformation
- Special case

MGF
Find MGF of $X_1 \ldots X_n$ to find MGF of Y

CDF Technique

Obj: get the CDF of Y from CDF of $X_1 \dots X_n$

① Discrete case:

General strategy:

1. $\{x_1 \dots x_n : h(x_1, \dots x_n) = y\}$ ⇒ Find this
   ↳ Across ALL Y.
2. $P((x_1, \dots, x_n))$
3. Piecewise function. → PMF
4. Convert to CDF : sum over PMF

Ex: ///

$$f_X(x) = \begin{cases} 1/4 & \text{if } |x| = 1 \\ 1/2 & \text{if } x = 0 \\ 0, & \text{o.w.} \end{cases}$$

$Y = X^2$. Find distribution of Y.

① Find the set:

a) What are possible values of Y:

$Y \in \{0, 1\}$

b) Find which values correspond to Y's output

$$S = \begin{cases} 1 & \text{if } x_n = -1, 1 \Rightarrow P(x = -1, 1) = 1/2 \\ 0 & \text{if } x_n = 0 \Rightarrow P(x = 0) = 1/2) \end{cases}$$

② Probabilities:

$$P(Y = y) = \begin{cases} 1/2 & \text{if } y = 0 \quad (P(X=0)) \\ 1/2 & \text{if } y = 1 \quad (P(X=1) + P(X=-1)) \\ 0 & \text{o.w.} \end{cases}$$

③ CDF.

$$P(Y \le y) = \begin{cases} 1/2 & y = 0 \\ 1 & y = 1 \end{cases}$$

② Continuous case:

① $\forall y \in \mathbb{R} \Rightarrow B_Y = \{(x_1, \dots x_n) : h(x_1, \dots x_n) \le y\}$

② Find CDF of Y:

$$F_Y(y) = P(Y \le y)$$
$$= P(h(x_1, \dots x_n) \le y)$$
$$= \int\int_{B_Y}\int f(x_1, \dots x_n)\, dx_1 \dots dx_n$$

③ Find PDF via $F_Y'(y)$

Ex://

$X \sim Exp(\theta)$

$\Downarrow$

$$F(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-x/\theta}, & x > 0 \end{cases}$$

$\Rightarrow$ $Y = F(X)$. Find distribution of Y.

① Understand function:

$$Y = F(X) = \begin{cases} 0, & x \le 0 \\ 1 - e^{x/\theta} & x > 0 \end{cases}$$

② Create set:

What values in set s.t. we can satisfy $F_Y(y)$?

Split into cases:

$y < 0$: when can $F(X) < 0 \Rightarrow \emptyset$ $\quad P(y \le 0) = 0$

$y = 0$: when can $F(X) = 0 \Rightarrow \{x : x \le 0\}$

$y \ge 1$: when can $F(X) \ge 1 \Rightarrow$ all $\mathbb{R}$

$0 < y < 1$: when can $0 < F(x) < 1$?

$\hookrightarrow \{x : F(x) \le y\} = \{x : x \le 0\} \cup \{1 - e^{x/\theta} \le y\}$

$= \{x : x \le 0\} \cup \{x : x \le -\theta \log(1-y)\}$

③ Create CDF:

$$F_Y(y) = P(X \in R_y) = \begin{cases} 0 & \text{if } y \le 0 \\ y & \text{if } 0 < y < 1 \\ 1 & \text{if } y \ge 1 \end{cases} \Rightarrow F(-\theta \log(1-y))$$

$P(X \le 0) = 0$

1-1 Transformation

- Univariate 1-1 Transformation Theorem:

X is a cts r.v w/ support A & $h(\cdot)$ is 1-1 on A

$$f(h(x)) = \begin{cases} f_X(h^{-1}(y)) \cdot \left| \frac{d}{dy} h^{-1}(y) \right|, & y \in \{h(x) : x \in A\} \\ 0, & \text{o.w.} \end{cases}$$

Useful for quick calculations.

Ex://

$$f(x) = \begin{cases} \frac{\theta}{x^{\theta+1}} & x > 1 \\ 0, & \text{o.w.} \end{cases}$$

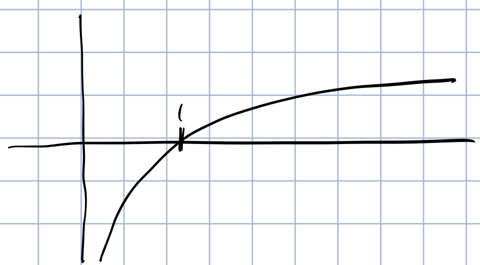
$Y = \log(X)$. Find distr of Y.

① Do we have 1-1?

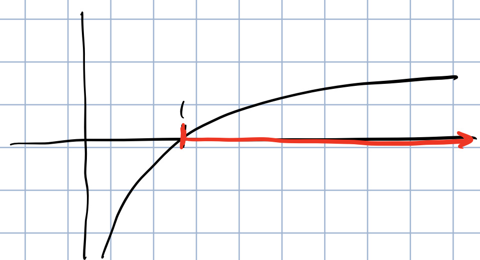
i) Look at X's support:

$A = (1, \infty)$

ii) Look at transformation



iii) Look if support is 1-1



② Find $h^{-1}(y)$:

$$y = h(x) = \log(x)$$

Solve for $x \Rightarrow \quad x = e^y$

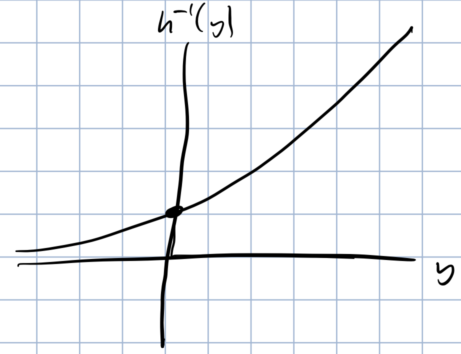
③ Find support of y:

$$y \in \{ h(x) : x \in A \}$$

$$= y \in \{ \log(x) : x > 1 \}$$

$$= y \in \{ y : e^y > 1 \}$$

$$= y \in \{ y : y > 0 \}$$



④ Find $\left| \frac{\partial}{\partial y} h^{-1}(y) \right|$ & $f_x (h^{-1}(y))$

$$\frac{\partial}{\partial y} h^{-1}(y) = e^y$$

$$f_x(h^{-1}(y)) = \frac{\theta}{(h^{-1}(y))^{\theta+1}} = \frac{\theta}{e^{y(\theta+1)}}$$

⑤ Combine everything:

$$f_Y(y) = \begin{cases} \frac{\theta}{e^{y(\theta+1)}} e^y , & y > 0 \\ 0. & \text{o.w.} \end{cases}$$

$$= \begin{cases} \frac{\theta}{e^{\theta+1}} , & y > 0 \\ 0, & \text{o.w.} \end{cases}$$

Bivariate Transformation:

Obj. X, Y, cts. r.v. $U = h_1(X, Y)$, $V = h_2(X, Y)$. Find distr. of U, V.

$U = h_1(X, Y)$, $V(X, Y)$ 1-1 transform on joint support of X, Y, then joint pdf of U, V is:

$$g(u,v) = \begin{cases} f_{X,Y}(w_1(u,v), w_2(u,v)) \cdot \left| \frac{\partial(w_1, w_2)}{\partial(x,y)} \right| , & (u,v) \in B \\ 0, & \text{o.w.} \end{cases}$$

Clarify:

$$w_1 = h_1^{-1}(u,v), \quad w_2 = h_2^{-1}(u,v)$$

$B = \{ (h_1(x,y), h_2(x,y)) : (x,y) \in A \}$ ⇒ Transform joint support.

Ex:// $\begin{pmatrix} X \\ Y \end{pmatrix} \sim BVN\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$. $U = X + Y$, $V = X - Y$. Find joint pdf of U, V.

① Determine dist. of variables under transform (x, y):

$$f_{X,Y}(x,y) = \frac{1}{2\pi} \exp\left(-\tfrac{1}{2}(x^2 + y^2)\right)$$

② Determine if inverse.

Strategy $\Rightarrow$ find inverse mapping. 1-1 transform.

Inverse mapping? $X = w_1(u,v)$, $Y = w_2(u,v)$

i) Write out transformation:

$$U = X + Y$$
$$V = X - Y$$

ii) Solve linear equation for X, Y

$$2X = U + V$$

$$X = \frac{U+V}{2} \Rightarrow w_1$$

$$\therefore Y = \frac{U-V}{2} \Rightarrow w_2$$

1-1 transform!

③ Support:

Initial support: $\mathbb{R}$

Transforms don't change support: $\mathbb{R}$

④ Jacobian:

$$\frac{\partial(w_1, w_2)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial w_1}{\partial u} & \frac{\partial w_1}{\partial v} \\ \frac{\partial w_2}{\partial u} & \frac{\partial w_2}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\tfrac{1}{4} - \tfrac{1}{4} = -\tfrac{1}{2}$$

$$\therefore \left| \frac{\partial(w_1, w_2)}{\partial(u,v)} \right| = \tfrac{1}{2}$$

⑤ Plug in inverse into joint pdf:

$$f_{X,Y}(w_1(u,v), w_2(u,v)) = \frac{1}{2\pi} \exp\left(-\tfrac{1}{2}\left[\left(\frac{u+v}{2}\right)^2 + \left(\frac{u-v}{2}\right)^2\right)\right)$$

$$= \frac{1}{2\pi} \exp\left(-\tfrac{1}{4}\left(u^2+v^2\right)\right)$$

⑥ Combine:

$$g(Y) = f_{X,Y}\left(w_1(u,v), w_2(u,v)\right) \cdot \left| \frac{\partial(w_1, w_2)}{\partial(u,v)} \right|$$

$$= \frac{1}{2\pi} \exp\left(-\tfrac{1}{4}\left(u^2+v^2\right)\right) \cdot \tfrac{1}{2}$$

$$= \frac{1}{4\pi} \exp\left(-\tfrac{1}{4}\left(u^2+v^2\right)\right)$$

$$\therefore \quad Y \sim BVN\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}2 & 0\\0 & 2\end{pmatrix}\right)$$

Ex:// $X, Y \sim Exp(1)$ (iid). Show that $X+Y \sim Gamma(2,1)$

Good trick: make our own $U, V \Rightarrow$ integrate out other variable.

$$U = X + Y$$
$$V = X \Rightarrow \text{will be integrate out}$$

① Understand what we're working with

$$f_{X,Y}(x,y) = \begin{cases} e^{-x} e^{-y} & 0 < x < \infty,\ 0 < y < \infty \\ 0 & \text{o.w.} \end{cases}$$

② Inverse check:

$$X = V$$
$$Y = U - V \quad \checkmark$$

③ Find our support:

$0 < x < \infty \Rightarrow 0 < v < \infty$

$0 < y < \infty \Rightarrow 0 < u - v < \infty$

$v < u < \infty$

$0 < u - v$

$v < u$

$u - v < \infty$

$u < \infty$

$$\therefore \{(u,v) : 0 < v < u < \infty\}$$

④ Find Jacobian:

$$\frac{\partial(w_1, w_2)}{\partial(u,v)} = \left|\begin{matrix} \frac{\partial w_1}{\partial u} & \frac{\partial w_2}{\partial u} \\ \frac{\partial w_2}{\partial u} & \frac{\partial w_2}{\partial v} \end{matrix}\right| = \left|\begin{matrix} 0 & 1 \\ 1 & -1 \end{matrix}\right| = -1 \Rightarrow 1 \text{ b/c abs.}$$

⑤ Joint support sub.

$$f_{X,Y}(w_1(u,v), w_2(u,v)) = \begin{cases} e^{-v} e^{-u+v} = e^{-u} & 0 < v < u < \infty \\ 0 & \text{o.w.} \end{cases}$$

⑥ Combine :

u still depends on v!

$$g_{U,V}(u,v) = \begin{cases} e^{-u} & 0 < v < u < \infty \\ 0 & \text{o.w.} \end{cases}$$

⑦ Integrate out v:

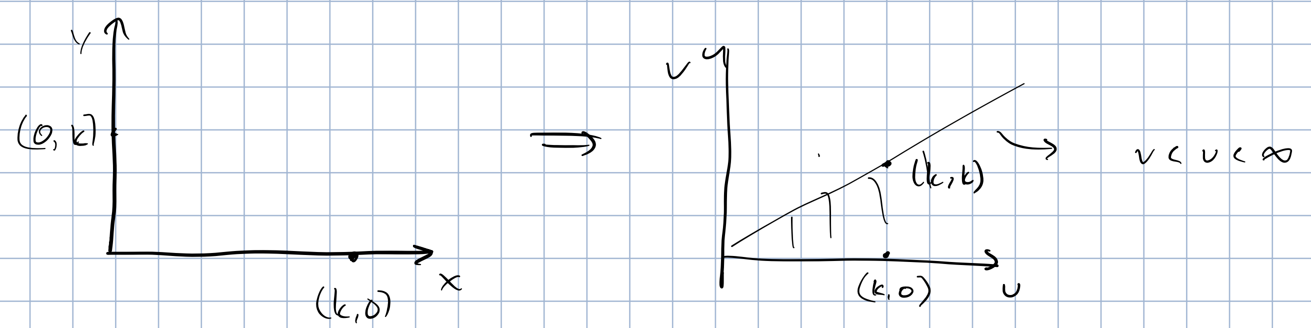
$$f_U(u) = \int_0^u e^{-u}\,dv = ve^{-u}\Big|_0^u = ue^{-u}$$

$$\therefore X+Y = \begin{cases} ue^{-u} & 0 < u < \infty \\ 0 & \text{o.w.} \end{cases}$$

$$= \text{Gamma}(2,1)$$

Small tip: How to find transformed support.

↳ Look at boundaries!



MGF Technique

$$X = Y \iff M_X(\vec{t}) = M_Y(\vec{t})$$

Steps:

① Find $M_Y(t) = E[e^{t h(x_1, x_2 \ldots)}]$

② Manipulate into MGF form that you recognize

③ Apply uniqueness theorem to find Y distn.

Ex:// $X \sim N(\mu, \sigma^2)$. $Y = aX + b$. $M_X(t) = \exp\{\mu t + \frac{t^2\sigma^2}{2}\}$.

Find dist. of Y.

① MGF of Y:

$$M_Y(t) = E[e^{t(aX+b)}]$$
$$= E[e^{taX} e^{tb}]$$
$$= e^{bt} \cdot M_X(at)$$
$$= e^{bt} \cdot e^{\mu a t + \frac{a^2t^2\sigma^2}{2}}$$
$$= e^{(\mu a + b)t + \frac{a^2t^2\sigma^2}{2}}$$

Exactly like MGF of $N(a\mu + b, a^2\sigma^2)$

② Uniqueness Theorem:

$$Y \sim N(a\mu + b, a^2\sigma^2)$$

Ex:// $X_1, \ldots, X_n$. $X_i \sim N(\mu_i, \sigma_i^2)$. $Y = \sum_{i=1}^{n} a_i X_i$.

Find Y.

① MGF of Y.

$$M_Y(t) = E[e^{t \sum_{i=1}^{n} a_i X_i}]$$
$$= \prod_{i=1}^{n} e^{t a_i X_i}$$
$$= \prod_{i=1}^{n} M_X(t a_i)$$
$$= \prod_{i=1}^{n} \exp\{\mu_i t a_i + \frac{t^2 a_i^2 \sigma_i^2}{2}\}$$
$$= \exp\{t \sum_{i=1}^{n} \mu_i a_i + \frac{t^2 \sum_{i=1}^{n} a_i^2 \sigma_i^2}{2}\}$$

② Uniqueness theorem:

$$M_Y(t) \sim N\left(\sum_{i=1}^{n} \mu_i a_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

$\therefore Y \sim$ "

Transformation on Standard Normal.

Distributions

χ^2 dist.

$$\boxed{\sum_{i=1}^{n} Z_i^2 \sim \chi^2_{(n)}}$$

t distr.

$Y \sim \chi^2_{(n)}$

$Z \perp Y$

$$\therefore \boxed{\frac{Z}{\sqrt{Y/n}} \sim t_{(n)}}$$

F distr.

$$\left.\begin{array}{l} Y_1 \sim \chi^2_{(n_1)} \\ Y_2 \sim \chi^2_{(n_2)} \end{array}\right\} Y_1 \perp Y_2$$

$$\boxed{\frac{Y_1/n_1}{Y_2/n_2} \sim F(n_1, n_2)}$$

Results

① MGF of $\chi^2_{(k)}$

$$M_X(t) = (1-2t)^{-k/2} \qquad \forall\, t \in \mathbb{R}$$

② Sum of χ^2

$$X_1 \ldots X_n, \quad X_i \sim \chi^2_{(k_i)} \Rightarrow \sum_{i=1}^{n} X_i \sim \chi^2_{\left(\sum_{i=1}^{n} k_i\right)}$$

③ Strong Lemma

Sample mean.

$$X_1 \ldots X_n \overset{iid}{\sim} N(\mu, \sigma^2). \quad \bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}, \quad s^2 \Rightarrow \text{sample var.}$$

$$\bar{X} \perp s^2$$

Theorem #1 : $X_1 \ldots X_n \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow \dfrac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

We know that :

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Proof:

$$W = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n} \left(\frac{(X_i - \bar{X}) + (\bar{X} - \mu)}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sigma}\right)^2 + \sum_{i=1}^{n} \left(\frac{\overline{x} - \mu}{\sigma}\right)^2 + 2\left(\frac{\overline{x} - \mu}{\sigma^2}\right) \underbrace{\sum_{i=1}^{n} (x_i - \overline{x})}_{0}$$

$$\therefore \sum_{i=1}^{n} (x_i - \overline{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x}$$

$$= n\overline{x} - n\overline{x}$$

$$= 0$$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\therefore \underbrace{\sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma}\right)^2}_{W_1} = \underbrace{\sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sigma}\right)^2}_{W_2} + \underbrace{\frac{n(\overline{x} - \mu)^2}{\sigma^2}}_{W_3}$$

W_1: $$\frac{x_i - \mu}{\sigma} \sim Z \Rightarrow U_1 \sim \chi^2_n$$

W_2: $$\frac{(n-1)s^2}{\sigma^2} = \frac{(n-1) \cdot \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2}{\sigma^2}$$

$$\sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sigma}\right)^2$$

W_3: $$\frac{n(\overline{x} - \mu)^2}{\sigma^2} = \left(\frac{\overline{x} - \mu}{\sigma / \sqrt{n}}\right)^2$$

$$\overline{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\frac{\overline{x} - \mu}{\sigma / \sqrt{n}} \sim Z$$

$$\left(\frac{\overline{x} - \mu}{\sigma / \sqrt{n}}\right)^2 \sim \chi^2_{(1)}$$

Use MGF technique:

$$W_1 = W_2 + W_3$$

$$E[e^{tW_1}] = E[e^{t(W_2 + W_3)}]$$

$$M_{W_1}(t) = M_{W_2}(t) M_{W_3}(t)$$

$$(1-2t)^{-n/2} = M_{W_2}(t) \cdot (1-2t)^{-1/2}$$

$$M_{W_2}(t) = (1-2t)^{-\frac{(n-1)}{2}}$$

$$\therefore \quad W_2 \sim \chi^2_{(n-1)}$$

Theorem #2: $X_1 \dots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. $\quad \dfrac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$

Proof:

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} = \frac{\frac{\bar{X}-\mu}{\sigma/n}}{\sqrt{s^2/\sigma^2}} = \frac{\frac{\bar{X}-\mu}{\sigma/n}}{\sqrt{\frac{\frac{(n-1)s^2}{\sigma^2}}{n-1}}}$$

$$f(\bar{x}) \quad \frac{\bar{X}-\mu}{\sigma/n} \Rightarrow Z \;, g(s^2) \quad \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$\therefore \quad f(\bar{x}) \perp g(s^2) \;!! \quad b/c \quad \bar{x} \perp s^2$$

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim \chi^2_{(n-1)}$$

Theorm #3: $X_1 \dots X_n \overset{iid}{\sim} N(\mu, \sigma_1^2)$. $\quad Y_1 \dots Y_n \overset{iid}{\sim} N(\mu, \sigma_2^2)$, $\quad X_i \perp Y_j$

then:

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$$

Proof:

$$\frac{(n_1-1)s_1^2}{\sigma_1^2} \sim \chi^2_{n-1}$$

$$\frac{(n_2-1)s_2^2}{\sigma_2^2} \sim \chi^2_{(n-1)}$$

$$\frac{\frac{\frac{(n_1-1)s_1^2}{\sigma_1^2}}{n_1-1}}{\frac{\frac{(n_2-1)s_2^2}{\sigma_2^2}}{n_2-1}} \qquad \begin{matrix} \to f(X_i) \\ \updownarrow \perp \\ \Rightarrow g(Y_i) \end{matrix} \qquad \therefore F(n_1-1, n_2-1)$$

CHAPTER 5: LIMITING / ASYMPTOTIC DISTR.

Motivation: distr. for lots of variables combined together.

$X_1, X_2, \ldots X_n \Rightarrow$ What is behavior of dist. as $n \to \infty$

Convergence

In distribution.

$X_n \xrightarrow{d} X$

Dist. of $X_n \approx$ Dist. of X as $n \to \infty$

In Prob.

$X_n \xrightarrow{P} X$

As $n \to \infty$, $P(X_n = x) = 1$

Conv. In Distr.

Defn: $X_1, X_2, \ldots, X_n$. X_n has CDF $F_n(x)$ $\forall n = 1, 2, \ldots$

X is a r.v. w/ CDF of $F(x)$.

$\lim_{n \to \infty} F_n(x) = F(x) \Rightarrow X_n \xrightarrow{d} X$

Note #1: $X_n \xrightarrow{d} X \not\Rightarrow X = X_n$

Ex: $W \sim$ Bernoulli $(1/2)$. $X = 1 - W$. $X_n = W$, $\forall n = 1, 2 \ldots$

① CDF of X_n:

$$F_n(x) = \begin{cases} 0 & 0 \leq x \\ 1/2 & 0 < x < 1 \\ 1 & 1 \geq x \end{cases}$$

② CDF of asymp. distr.

$$F(x) = \begin{cases} 0 & \_\_\_ \\ 1/2 & \_\_\_ \\ 1 & \_\_\_ \end{cases}$$

③ Limit of $F_n(x)$:

$$\lim_{n \to \infty} F_n(x) = F(x) \; !!$$

However, $X_n \neq X$! X_n is off by 1.

Note #2: Only care about cont. parts of r.v.

Ex:// $X_n \sim \text{Unif}(0, 1/n)$. $X = 0$.

Show $X_n \xrightarrow{d} X$

① CDF $F_n(x)$

$$F_n(x) = \begin{cases} 0, & x \le 0 \\ nx, & 0 < x < 1/n \\ 1, & x \ge 1/n \end{cases}$$

② CDF of limiting dist.

$$F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \ge 0 \end{cases}$$ ⇒ Point of discontinuity @ $x = 0$.

③ $\lim_{n\to\infty}$

$$\lim_{n\to\infty} F_n(x) = \begin{cases} 0, & x \le 0 \\ 1, & x > 0 \end{cases}$$

$X_n \xrightarrow{d} X$! Even if there is discontinuities.

Tip:

$$\lim_{n\to\infty} \left[1 + \frac{b}{n} + \frac{\psi(n)}{n}\right]^{cn} = e^{bc}$$

Ex:// $X_1, \dots, X_n \overset{iid}{\sim} \text{Unif}(0,1)$. $X_{(1)} = \min\{X_1, \dots, X_n\}$ $X_{(n)} = \max\{X_1, \dots X_n\}$

a) Find asymptotic dist. of $nX_{(1)}$.

① Find CDF of $nX_{(1)}$

Recall $\min\{X_1, \dots, X_n\}$

↳ $$F_{X_{(1)}}(y) = \begin{cases} 0, & y \le 0 \\ 1-(1-y)^n, & 0 \le y < 1 \\ 1, & y > 1 \end{cases}$$

Use CDF method to find dist.

$$\begin{aligned} F_{nX_{(1)}}(x) &= P(nX_{(1)} \le x) \\ &= P\left(X_{(1)} \le \frac{x}{n}\right) \\ &= F_{X_{(1)}}(x/n) \end{aligned}$$

$$= \begin{cases} 0, & x/n \le 0 \\ 1-(1-x/n)^n, & 0 \le x/n < 1 \\ 1, & x/n > 1 \end{cases}$$

$$= \begin{cases} 0, & x \le 0 \\ 1-(1-x/n)^n, & 0 \le x < n \\ 1, & x > n \end{cases}$$

② Limit

$$\lim_{n\to\infty} F_{nX_{(1)}}(x) = \begin{cases} 0, & x \le 0 \\ \lim_{n\to\infty} 1-(1-x/n)^n, & 0 < x < \infty \end{cases}$$

$$1-\Big(1+\frac{\overset{b}{(-1)}\cdot x}{n}\Big)^{\overset{c=1}{n}} = e^{-x}$$

$$= \begin{cases} 0, & x \le 0 \\ 1-e^{-x}, & 0 < x < \infty \end{cases}$$

$$\therefore \quad nX_{(1)} \xrightarrow{d} \text{Exp}(1)$$

Convergence in Probability.

Defn: $X_1, \ldots, X_n$, X. $X_n \xrightarrow{P} X \Leftrightarrow \forall\, \varepsilon > 0$, $\lim_{n\to\infty} P(|X_n - X| \ge \varepsilon) = 0$

↳ Informal: prob. of X_n differing from X is really small as $n \to \infty$

↳ Equivalent: $\lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$

$\lim_{n\to\infty} P(X_n \in (X-\varepsilon, X+\varepsilon)) = 1$

} Virtually approaching same value.

Ex:// $W \sim \text{Unif}(0,1)$.

$$X_n = \begin{cases} 1, & 0 < w < 1/n \\ 0, & \text{o.w.} \end{cases}$$

$$X = 0$$

} Show $X_n \xrightarrow{P} X$

① Define ε:

Let $\varepsilon > 0$

② Un probability:

$$P(|X_n - X| \geq \varepsilon) = P(X_n \geq \varepsilon)$$
$$= P(X_n = 1) \Rightarrow X_n \geq \varepsilon \text{ if } X_n = 1$$
$$= P(0 < w < 1/n)$$
$$= \int_0^{1/n} dx = 1/n$$

We want prob. in terms of n:

$$P(|X_n - X| \geq \varepsilon) = f(n)$$

③ Limit of prob.

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = \lim_{n\to\infty} 1/n = 0$$

$\therefore$ By defn. $X_n \xrightarrow{P} X$

Ex: $X_n \sim$ Bernoulli $(1 - 1/n)$. $X = 1$. Show $X_n \xrightarrow{P} X$.

Assume $\varepsilon > 0$.

① Find $P(|X_n - X| \geq \varepsilon) = f(n)$

$$P(|X_n - X| \geq \varepsilon) = P(|X_n - 1| \geq \varepsilon)$$

$$|X_n - 1| = \begin{cases} 1 & \text{if } X_n = 0 \\ 0 & \text{if } X_n = 1 \end{cases}$$

$\Rightarrow$ May have to split into piecewise & figure when $|X_n - 1| > 0$

$$P(|X_n - 1| \geq \varepsilon) = P(|X_n - 1| = 1)$$
$$= P(X_n = 0)$$
$$= 1 - (1 - 1/n)$$
$$= 1/n$$

By Bernoulli

② Limit:

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = \lim_{n\to\infty} 1/n = 0$$

Theorem: $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

- Useful if $X_n \xrightarrow{d} X$ is difficult $\Rightarrow$ just compute $X_n \xrightarrow{P} X$
- Converse is not true
  - ↳ Exception: $\forall c \in \mathbb{R}$, $X_n \xrightarrow{d} c \Rightarrow X_n \xrightarrow{P} c$

Proof: Suppose $X_n \xrightarrow{d} c$, $\varepsilon > 0$.

$$\lim_{n \to \infty} P(|X_n - X| \geq \varepsilon) = 0$$

Useful trick: squeeze theorem. Show that upper bound $= 0$. as $n \to \infty$

$$\begin{aligned}
P(|X_n - X| \geq \varepsilon) &= P(|X_n - c| \geq \varepsilon) \\
&= P(X_n \geq c + \varepsilon) + P(X_n \leq c - \varepsilon) \\
&= 1 - P(X_n \leq c + \varepsilon) + F_n(c - \varepsilon) \\
&\leq 1 - P(X_n \leq c + \tfrac{\varepsilon}{2}) + F_n(c - \varepsilon) \\
&\quad 1 - F_n(c + \varepsilon/2) + F_n(c - \varepsilon)
\end{aligned}$$

$$F(x) = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } x \geq c \end{cases} \qquad \lim_{n \to \infty} F_n(x) = F(x) \text{ by conv. in distr.}$$

$$\begin{aligned}
\lim_{n \to \infty} \left(1 - F_n(c + \varepsilon/2) + F_n(c - \varepsilon)\right) &= 1 - F(c + \varepsilon/2) + F(c - \varepsilon) \\
&= 1 - 1 + 0 \\
&= 0
\end{aligned}$$

By squeeze theorem: $P(|X_n - X| \geq \varepsilon) = 0$

Squeeze theorem is useful. We want to take upper bounds easily

↳ Markov's Inequality:

X is a r.v. $\forall k > 0, c \geq 0$ $\quad P(|X| \geq c) \leq \dfrac{E[|X|^k]}{c}$

Usually, $k = 2$ is nice $\Rightarrow$ we can derive $E[|X|^2]$ easily

↳ Chebyshev's Inequality

$$P(|X - E[X]| \geq k\sqrt{Var(X)}) \leq \frac{1}{k^2}$$

Ex:// $Y \sim \text{Unif}(0,1)$. $X_n = Y^n$. Show $X_n \xrightarrow{p} X$ if $X=0$.

Let $\varepsilon > 0$.

① Find $P(|X_n - X| > \varepsilon) = f(n)$

$$P(|X_n - X| > \varepsilon) = P(|Y^n| > \varepsilon)$$

② If too difficult $\Rightarrow$ squeeze theorem w/ inequalities.

By Markov's Ineq.

$$P(|X_n - X| \geq \varepsilon) = \frac{E[|X_n - X|]}{\varepsilon}$$

$$= \frac{E[X_n]}{\varepsilon}$$

$\therefore$ Find $E[X_n]$

$$E[X_n] = E[Y^n] = \int_0^1 y^n \, dy$$

$$= \frac{y^{n+1}}{n+1} \Big|_0^1$$

$$= \frac{1}{n+1}$$

$$\therefore P(|X_n - X| \geq \varepsilon) \leq \frac{1}{\varepsilon(n+1)}$$

$\lim_{n \to \infty} \frac{1}{\varepsilon(n+1)} = 0 \Longrightarrow$ By squeeze theorem

↳ $X_n \xrightarrow{p} X$

Result: $X_1, \ldots,$ w/ $E[X_i] = \mu$, $\text{Var}[X_n] = \sigma^2$, $\lim_{n \to \infty} \sigma^2 = 0$

↳ $X_n \xrightarrow{p} \mu$

Weak Law of Large Numbers: $X_1, \ldots \overset{iid}{\sim} E[X_i] = \mu$, $\text{Var}[X_i] = \sigma^2 \Rightarrow \overline{X}_n \xrightarrow{p} \mu$ $\sum_{i=1}^{n} \frac{X_i}{n}$

Ex:// $X_1 \dots X_n \overset{iid}{\sim}$ Bernoulli (p). $W_n = \sum_{i=1}^{n} \frac{2^{x_i}}{n}$. W_n convergence?

Trick: to use WLLN $\Rightarrow$ convert into the form $\frac{1}{n} \sum \frac{Y_i}{n}$

① Variable transform:

$$Y_i = 2^{x_i}$$

② Mean, variance

$$E[Y_i] = E[2^{x_i}]$$

$$= 2(p) + 1(1-p)$$

$$= 2p + 1 - p$$

$$= 1 + p$$ $\Rightarrow$ All Y_i have same mean & variance

$$E[Y_i^2] = E[2^{2x_i}]$$

$$= 4(p) + 1(1-p)$$

$$= 4p + 1 - p$$

$$= 1 + 3p$$

$$Var[Y_i] = E[Y_i^2] - E[Y_i]^2 < \infty$$

③ Apply WLLN:

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \xrightarrow{P} 1 + p$$

Central Limit Theorem

Theorem: $X_1 \dots X_n$ iid. $E[X_i] = \mu$, $Var[X_i] = \sigma^2 < \infty$

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$$

We don't need to be using $\bar{X}_n$ to apply CLT. This can also apply to sum of r.v.

Ex:// $X_n \sim \chi^2_{(n)}$. Show $\frac{X_n - n}{\sqrt{2n}} \xrightarrow{d} N(0,1)$

① Find sum:

$$X_n = \sum_{i=1}^{n} \chi^2_{(1)}$$

② Apply CLT: all X_i have same mean & finite variance

We need to calculate this parameters knowing $X_i = \chi^2_{(1)}$

a) $\chi^2_{(1)}$:

$$E[\chi^2_{(1)}] = 1$$

$$Var[\chi^2_{(1)}] = \underbrace{E[(\chi^2_{(1)})^2]}_{??} - 1^2 =$$

Use MGF:

$$M_{\chi^2_{(1)}}(t) = (1-2t)^{-1/2}$$

$$M'_{\chi^2_{(1)}}(t) = -\frac{1}{2}(1-2t)^{-3/2}(-2) = (1-2t)^{-3/2}$$

$$M''_{\chi^2_{(1)}}(t) = 3(1-2t)^{-5/2}$$

$$M''_{\chi^2_{(1)}}(0) = 3(1)^{-5/2} = 3$$

$$\therefore Var[\chi^2_{(1)}] = 3 - 1 = 2$$

③ CLT:

$$\frac{\sqrt{n}\left(\frac{\overbrace{\sum X_i}^{X_n}}{n} - 1\right)}{\sqrt{2}} = \frac{\frac{X_n}{\sqrt{n}} - \sqrt{n}}{\sqrt{2}}$$

$$= \frac{X_n - n}{\sqrt{2n}} \xrightarrow{d} N(0,1) \text{ by CLT}$$

Ex:// $X_1 \dots X_n \overset{iid}{\sim} Unif(0,1)$. Show that $\dfrac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^3 - 1/4\right)}{\sqrt{\frac{1}{7} - (1/4)^2}} \xrightarrow{d} N(0,1)$

① Sum: $\sum_{i=1}^{n} \boxed{X_i^3} \Rightarrow$ Find $E[X_i^3]$ & $Var[X_i] < \infty$

② Verify CLT app.: for $E[X_i^3]$ & $Var[X_i]$.

$$E[X_i^3] = \int_0^1 x^3\,dx = 1/4$$

$$E\left[(X_i^3)^2\right] = \int_0^1 x^6\,dx = \frac{1}{7}$$

$$Var\left[X_i^3\right] = E\left[(X_i^3)^2\right] - E\left[X_i^3\right]^2$$
$$= 1/7 - (1/4)^2$$

③ Apply CLT:

$$\frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^3 - 1/4\right)}{\sqrt{\frac{1}{7} - (1/4)^2}} \xrightarrow{d} N(0,1)$$

Limit Theorems

1. Continuous Mapping Theorem: Assume cont. func. $g(\cdot)$

   i) $\forall c \in \mathbb{R}$ if $X_n \xrightarrow{p} c \Rightarrow g(X_n) \xrightarrow{p} g(c)$

   ii) $X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X)$

2. Slutsky's Theorem: $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{p} c$

   i) $X_n + Y_n \xrightarrow{d} X + c$

   ii) $X_n Y_n \xrightarrow{d} cX$

   iii) $\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c}$ $(c \neq 0)$

Ex:// $X_1, \ldots X_n \overset{iid}{\sim}$ Poisson (λ). $\mu_n = \sqrt{n}\,(\bar{X}_n - \lambda)$. $Z_n = \frac{\mu_n}{\sqrt{\bar{X}_n}}$. Find convergence of μ_n & Z_n.

a) Convergence of μ_n

① Looks like CLT

$$\frac{\sqrt{n}\,(\bar{X}_n - \lambda)}{\sqrt{\lambda}} \xrightarrow{d} N(0,1) \text{ by CLT}$$

② Use Limit Theorem:

$$A = \frac{\sqrt{n}\,(\bar{X}_n - \lambda)}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$$

$$c = \sqrt{\lambda} \xrightarrow{p} \sqrt{\lambda}$$

$\therefore$ By Slutsky's theorem,

$$cA = \sqrt{n}(\bar{X}_n - \lambda) \xrightarrow{d} \sqrt{\lambda}\, N(0,1)$$
$$\xrightarrow{d} N(0, \lambda)$$ By prop. of N

b) Z_n:

$$Z_n = \frac{M_n}{\sqrt{\bar{X}_n}} \Rightarrow \text{We need } \sqrt{\bar{X}_n} \xrightarrow{p} c$$

① Show $\sqrt{\bar{X}_n} \xrightarrow{p} \dots$

General trick: if using $\bar{X}_n$ exclusively $\Rightarrow$ WLLN!

By WLLN: $\bar{X}_n \xrightarrow{p} \lambda$

By Cont. Mapping Theorem: $\sqrt{\bar{X}_n} \xrightarrow{p} \sqrt{\lambda}$

By Slutsky's Theorem: $\frac{M_n}{\sqrt{\bar{X}_n}} \xrightarrow{d} \frac{N(0,\lambda)}{\sqrt{\lambda}} = N(0,1)$

Ex.// $X_1 \dots X_n \overset{iid}{\sim} \text{Unif}(0,1)$. $U_n = \max_{1 \le i \le n} X_i$. $V_n = e^{-n(1-U_n)}$. Find asymp. distr. of V_n.

① Convergence in distribution.

$$n(1-U_n) \xrightarrow{d} \text{Exp}(1)$$

② Continuous mapping theorem.

$$g(\underbrace{n(1-U_n)}) \xrightarrow{d} g(\overset{X}{\text{Exp}(1)})$$

$$e^{-(\ d\ )} \xrightarrow{d} e^{-X}$$

③ Transformation techniques.

Let $T = e^{-Y}$. By CDF tech.

$$F_T(t) = P(e^{-Y} \le t)$$

$$\vdots$$

$$\begin{cases} 0, & t \le 0 \\ t, & 0 < t < 1 \\ 1, & t \ge 1 \end{cases}$$

$\therefore T \sim \text{Unif}(0,1)$

$\therefore V_n \xrightarrow{d} \text{Unif}(0,1)$

3. Delta Method

$X_1 \dots X_n$ r.v. s.t. $a_n(X_n - \theta) \xrightarrow{d} N(0, \sigma^2)$ and $\lim_{n\to\infty} a_n = \infty$

and $g(x)$ is diff. at $x = \theta$ and $g'(\theta) \neq 0$

Then: $a_n(g(X_n) - g(\theta)) \xrightarrow{d} N(0, (g'(\theta))^2 \sigma^2)$

Tip: $a_n(X_n - \theta) \xrightarrow{d} N(0, \sigma^2)$ often by CLT

When to use Delta Method:

- Looks really similar to CLT
- Function of $\bar{X}_n$ & means

Ex:// $X_1 \dots \overset{iid}{\sim} Poi(\lambda)$. Find limiting distn. of $\sqrt{n}(\sqrt{\bar{X}_n} - \sqrt{\lambda})$

① CLT

$$\frac{\sqrt{n}(\bar{X}_n - \lambda)}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$$

$$\sqrt{n}(\bar{X}_n - \lambda) \xrightarrow{d} N(0, \lambda)$$

② Find $g(\cdot)$ to use Delta Method.

$$g(t) = \sqrt{t}$$

③ Check if conditions satisfied.

$$\lim_{n\to\infty} \sqrt{n} = \infty \quad \checkmark$$

$$g'(\lambda) = (\sqrt{\lambda})' = \frac{1}{2\sqrt{\lambda}} \neq 0$$

④ Apply Delta Method.

$$\sqrt{n}(\sqrt{\bar{X}_n} - \sqrt{\lambda}) \xrightarrow{d} N\left(0, \left(\frac{1}{2\sqrt{\lambda}}\right)^2 \lambda\right)$$

$$\xrightarrow{d} N(0, 1/4)$$

Ex:// $X_1 \dots$ w/ mean 0 and variance $\sigma^2 < \infty$. Find distn. of $\bar{X}_n^2$

① CLT:

$$\frac{\sqrt{n}(\bar{X}_n - 0)}{\sigma} \xrightarrow{d} N(0,1)$$

$$\sqrt{n}\,(\bar{X}_n) \xrightarrow{d} N(0, \sigma^2)$$

② Continuogous function:

$$g(t) = t^2$$

③ Condition check:

$$g(0) = (0^2)' = 0 \quad \times$$

Continuous mapping on CLT:

$$\frac{\sqrt{n}\,\bar{X}_n}{\sigma} \xrightarrow{d} N(0,1)$$

$$\left(\frac{\sqrt{n}\,\bar{X}_n}{\sigma}\right)^2 \xrightarrow{d} \chi^2_{(1)}$$

$$n\bar{X}_n^2 \xrightarrow{d} \sigma^2 \chi^2_{(1)}$$

$$\bar{X}_n^2 \xrightarrow{d} \frac{\sigma^2}{n} \chi^2_{(1)}$$

Summary:

① Building Blocks:

1. $\xrightarrow{d}$
   - I. By definition: $F_n(x) \longrightarrow F(x)$
   - II. CLT: sum of iid r.v.

2. $\xrightarrow{p}$
   - I. By definition: $\forall \varepsilon > 0, \quad \lim_{n \to \infty} P(|X_n - X| > \varepsilon) = 0$
     - ↳ Markov's Inequality
   - II. WLLN: sum of iid. r.v.

③ Glue:

1. Continous mapping: applied to entire block
   - Slutsky's: 2 r.v., one converging in prob.
2. Delta Method: $g(\cdot)$ continuous func. inside $\sqrt{n}(\ldots)$

Chapter 6: Point Estimation

Intro

Population variables : $X_1 \dots X_n \overset{iid}{\sim} f(x, \theta)$ → Parameters $\in \Omega$ (parameter space).

Observations : $x_1 \dots x_n$

Goal: given our observation ⇒ find good guess θ

Statistic : $T(\vec{X})$ (function of $\vec{X}$, not related to θ)

$$T(\vec{x}_0) = \frac{X_1 + \dots X_n}{n} \qquad T(\vec{X}) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)$$

◦ Estimator of $\vec{\theta}$ ⇒ $\hat{\theta}$

Method of Moments

Defn: $X_1 \dots X_n$ iid ~ pmf/pdf of $f(\vec{x}; \vec{\theta})$.

From sample ←
$$\hat{\mu}_1 = E[X_i] = g_1(\vec{\theta})$$
$$\hat{\mu}_2 = E[X_i^2] = g_2(\vec{\theta})$$
$$\vdots$$
From PDF/PMF

p times for p-dimension.

Method of Moments estimator (MME): solution to above for $\vec{\theta}$

Ex: // $X_1 \dots X_n \overset{iid}{\sim} Poi(\lambda)$. Find MME for λ.

1. First moment from sample:
$$E[X_i] = \sum_{i=1}^{n} \frac{X_i}{n}$$

2. First moment from distn.
$$E[X_i] = \lambda$$

3. Equate:
$$\hat{\lambda} = \sum_{i=1}^{n} \frac{x_i}{n}$$ ⇒ Function of random variables.

To judge quality:

1. Unbiased: $E[\text{estimator}] = \text{parameter}$.

② Consistent:

$$\text{estimator} \xrightarrow{P} \text{actual}$$

Ex:// $X_1 \dots X_n \overset{iid}{\sim}$ Unif $(0, \theta)$. Find MME for θ.

① Calculate 1^{st} moment from sample:

$$E[x_i] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

② Calculate 1^{st} moment from distribution:

$$E[x_i] = \frac{\theta}{2}$$

③ Equate:

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{\theta}{2}$$

$$\hat{\theta} = 2\bar{X}_n$$

④ Check if unbiased / consistent:

Unbiased:

$$E[\hat{\theta}] = \frac{2}{n}\sum_{i=1}^{n} E[x_i]$$

$$= \frac{2}{n}\sum_{i=1}^{n} \frac{\theta}{2}$$

$$= \theta \checkmark$$

Consistent:

By WLLN: $\bar{x}_n \xrightarrow{P} \frac{\theta}{2}$

By CMT:

$$2\bar{x}_n \xrightarrow{P} \theta \checkmark$$

Ex:// $X_1 \dots X_n \overset{iid}{\sim}$ Uniform $(-\theta, \theta)$. $\theta > 0$

① 1^{st} moment: empirically & theoretically

$$E[X_1] = E[X_1]$$

$$\bar{X}_n = \frac{\theta - \theta}{2} = 0 \dots$$

Can't do anything $\Rightarrow$ 2^{nd} moment.

② 2^{nd} moment:

$$E[X_1^2] = E[X_i^2]$$

$$E[X_1^2] = Var[X_1] + E[X_1]^2$$

$$= \frac{(\theta - (-\theta))^2}{12} + 0^2$$

$$= \frac{\theta^2}{3}$$

$$\therefore \quad \frac{1}{n}\sum_{i=0}^{n} x_i^2 = \frac{\hat{\theta}^2}{3}$$

$$\hat{\theta} = \sqrt{\frac{3\sum_{i=0}^{n} x_i}{n}}$$

You can use any moment to figure out estimate

↳ Convention: use first p moments

Ex:// $X_1, \ldots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

① Moments:

$$E[X_1] = \mu$$

$$E[X_1^2] = Var[X_1] + E[X_1]^2$$

$$= \sigma^2 + \mu^2$$

② Empirical moments:

$$\hat{\mu} = \frac{1}{n}\sum_{i=0}^{n} X_i$$

$$\hat{\sigma}^2 + \hat{\mu}^2 = \frac{1}{n}\sum_{i=0}^{n} X_i^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=0}^{n} X_i^2 - \overline{X}_n^2$$

② Check unbiased & consistent:

$$E[\hat{\sigma}^2] = \frac{1}{n}\sum_{i=0}^{n} E[X_i^2] - E[\overline{X}_n^2]$$

$$= E[X_i^2] - (Var[\overline{X}_n] + E[\overline{X}_n]^2)$$

$$= \sigma^2 + \mu^2 - \left(\frac{\sigma^2}{n} + \mu^2\right)$$

$$\neq \sigma^2 \Rightarrow \text{Biased!}$$

Maximum Likelihood

Idea: find param. that maximize probability that observations occur.

Likelihood function:

$$\mathcal{L}(\vec{\theta};\vec{x}) = f(\theta; x_1)\cdot f(\theta; x_2)\cdot \ldots$$

$$= \prod_{i=1}^{n} f(\vec{\theta};\vec{x})$$

To maximize:

① If θ is scalar:

$$\frac{d}{d\theta}\mathcal{L}(\hat{\theta}; x) = 0 \quad \Rightarrow \text{Which } \hat{\theta} \text{ gives us solution.}$$

② If θ is vector:

$$\nabla \mathcal{L}(\vec{\theta}; x) = \vec{0}$$

$$\begin{bmatrix} \frac{\partial}{\partial\theta_1}\mathcal{L}(\vec{\theta}; x) \\ \vdots \\ \frac{\partial}{\partial\theta_p}\mathcal{L}(\vec{\theta}; x) \end{bmatrix} = \vec{0}$$

③ Check boundary conditions of $\hat{\theta}$ to maximize.

Trick: log likelihood function

$$\ell(\vec{\theta}; x) = \log(\mathcal{L}(\vec{\theta}; \vec{x}))$$

$$= \log\left(\prod_{i=1}^{n} f(\vec{\theta}; x)\right)$$

$$= \sum_{i=1}^{n} \log f(\vec{\theta}; x)$$

The max of $\ell(\vec{\theta}; x)$ = max of $\mathcal{L}(\vec{\theta}; x)$

Ex:// $X_1 \ldots X_n \sim Poi(\lambda)$. Find MLE of λ

① Likelihood function:

$$\mathcal{L}(\lambda; x) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

② Log likelihood function:

$$\ell(\lambda; x) = \left(\sum_{i=1}^{n} x_i\right) \log(\lambda) + n \log(e^{-\lambda}) - \log\left(\prod_{i=1}^{n} x_i!\right)$$

$$= \left(\sum_{i=1}^{n} x_i\right) \log(\lambda) - n\lambda - \log\left(\prod_{i=1}^{n} x_i!\right)$$

③ Maximize:

$$\frac{\partial}{\partial \lambda} \ell(\lambda; x) = \frac{1}{\lambda}\left(\sum_{i=1}^{n} x_i\right) - n = 0$$

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{X}_n$$

Ex:// $X_1, \ldots, X_n \overset{iid}{\sim}$ Uniform $(0, \theta)$.

$$f(x; \theta) = \begin{cases} \frac{1}{\theta}, & 0 \le x_i \le \theta \\ 0, & \text{o.w.} \end{cases}$$

Find MLE for θ

① Likelihood function:

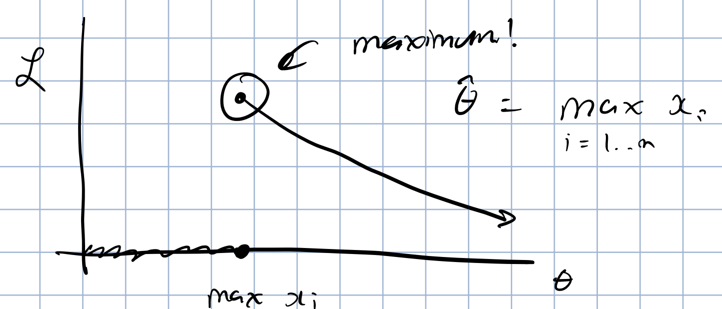
$$\mathcal{L}(\theta; x) = \prod_{i=1}^{n} f(\theta; x)$$

$$= \prod_{i=1}^{n} \begin{cases} 1/\theta, & 0 \le x_i \le \theta \\ 0, & \text{o.w.} \end{cases}$$

$$= \begin{cases} \theta^{-n}, & 0 \le \text{all } x_i \le \theta \\ 0, & \text{o.w.} \end{cases}$$

$$= \begin{cases} \theta^{-n} & \max x_i \le \theta \\ 0 & \text{o.w.} \end{cases}$$

② Graph:



If param in range ⇒ graph it out!

Theorem #1 (Invariance):

If $\hat{\theta}$ is MLE of θ, then $g(\hat{\theta})$ is MLE for $g(\theta)$.

Ex: // $X_1, \ldots X_n \overset{iid}{\sim} Poi(\lambda)$.

Known: $\hat{\lambda} = \overline{X}_n$. $E[X_i] = Var[X_i] = \lambda$

a) MLE for $\lambda(\lambda+1)$?

$$\widehat{(\lambda(\lambda+1))} = \overline{X}_n(\overline{X}_n + 1)$$

b) MLE for $P(X_i = 0)$

$$P(X_i = 0) = e^{-\lambda}$$

$$MLE \Rightarrow e^{-\overline{X}_n}$$

c) MLE for:

$$\mathbb{I} = \begin{cases} 1, & \lambda \geq 10 \\ 0, & o.w. \end{cases}$$

$$MLE \text{ for } \mathbb{I} = \begin{cases} 1 & \overline{X}_n \geq 10 \\ 0 & o.w. \end{cases}$$

Assume θ is scalar

Terminologies:

- Score function: $s(\theta; x) = \frac{d}{d\theta} \ell(\theta; x)$
- Information function: $I(\theta; \vec{x}) = -\frac{d}{d\theta^2} \ell(\theta; \vec{x})$
- Expected info function: $J(\theta) = E[I(\theta; \vec{x})]$

Theorem #2

Under regularity cond. $(\hat{\theta} - \theta)[J(\theta)]^{1/2} \xrightarrow{d} N(0,1)$.

Super powerful along w/ delta method. Can help us find distn. of quantities that depend on MLE

Ex: // Find the distribution of $P(X_i = 0)$ if $X_i \sim Poi(\lambda)$.

We know MLE of $P(X_i = 0) = e^{-\overline{X}_n}$

① Distn of MLE $\hat{\lambda}$ using prev. theorem.

$$\ell(\lambda; x) = \prod_{i=1}^{n} f(\lambda; x) = \left(\sum_{i=1}^{n} x_i\right) \log \lambda - n\lambda - \left(\sum_{i=1}^{n} \log(x_i!)\right)$$

$$S(\lambda; x) = \frac{1}{\lambda} \sum_{i=1}^{n} x_i - n$$

$$I(\lambda; x) = -\left(-\frac{1}{\lambda^2} \sum_{i=1}^{n} x_i\right) = \frac{1}{\lambda^2} \sum_{i=1}^{n} x_i$$

$$J(\theta) = \mathbb{E}\left[\frac{1}{\lambda^2} \sum_{i=1}^{n} X_i\right]$$

$$= \frac{1}{\lambda^2} \sum_{i=1}^{n} \mathbb{E}[X_i]$$

$$= \frac{1}{\lambda^2} n\lambda$$

$$= \frac{n}{\lambda}$$

From theorem:

$$\frac{\sqrt{n}(\overline{X}_n - \lambda)}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$$

② Use delta method to get estimate for $e^{-\overline{X}_n}$

$$g(t) = e^{-t}$$

$$g'(t) = -e^{-t} \Rightarrow g'(\lambda) = -e^{-\lambda} \neq 0$$

∴ By cts. mapping theorem:

$$\sqrt{n}(\overline{X}_n - \lambda) \xrightarrow{d} N(0, \lambda)$$

By delta method.

$$\sqrt{n}(e^{-\overline{X}_n} - e^{-\lambda}) \xrightarrow{d} N(0, \lambda e^{-2\lambda})$$

Solve for $e^{-\overline{X}_n}$

$$e^{-\overline{X}_n} - e^{-\lambda} \xrightarrow{d} N\left(0, \frac{\lambda e^{-2\lambda}}{n}\right)$$

$$e^{-\overline{X}_n} \xrightarrow{d} N\left(e^{-\lambda}, \frac{\lambda e^{-2\lambda}}{n}\right)$$

If finding result from MLE distn → use above theorem #2 + delta method.