Lecture 1

1. sample space, event, σ-field, probability, prob. space.

2. Conditional prob., independence, Bayes rule.

3. Random variable. $\{\Omega, \mathcal{F}\} \setminus \{\mathbb{R}, \mathcal{B}\}$.
measurable.

Mass function, CDF, pdf.

Normal: $\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(x-\mu)^2)$.

Poisson. $P(X=k) = e^{-\lambda} \lambda^k / k!$

Uniform
Bernoulli $\begin{cases} p \\ 1-p \end{cases}$
mean p.
var. $p(1-p)$
χ^2, t, ~~F~~

4. α-th quantile $\arg\min_{q \in \mathbb{R}} Pr(X \leq q) \geq \alpha$

Bayes rule

---

5. Multivariate $\{\Omega, \mathcal{F}\} \setminus \{\mathbb{R}^n, \mathcal{B}\}$.

$Pr(B) = \int_B f(x_1 \dots x_n) dx_1 \dots dx_n$

$\text{cdf}(x_1 \dots x_n) = \int_{\infty}^{x_1} \dots \int_{\infty}^{x_n} f(x_1 \dots x_n) dx_1 \dots x_n$

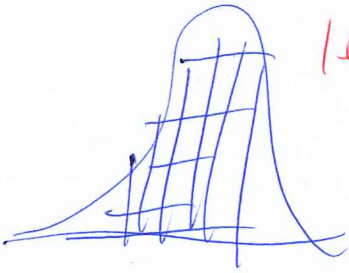
Bivariate: marginal: $f(x_1) = \int f(x_1, x_2) dx_2$.

Conditional $f(x_1 | x_2) = f(x_1, x_2) / f(x_1)$.

Multivariate normal: $\frac{1}{(\sqrt{2\pi})^n} |\Sigma|^{-\frac{1}{2}} \exp(-\frac{1}{2}(x-\mu)' \Sigma^{-1} (x-\mu))$

indep.

$c'X$. linearity.

Expectation $E(X) = \int x f(x)\, dx$.  lecture 2

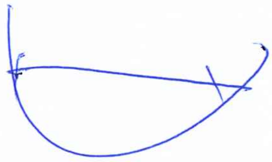
Bayes rule

$E(X) = \sum_{X=k} k\, P(X=k)$.

~~Func~~ Expectation of a function. $\int g(x) f(x)\, dx$

$= \int g(x)\, dF(x)$.

~~Pr~~ linearity: $E(a+bX) = a + bE(X)$.

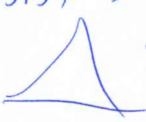
Jensen's.  if ϕ is convex, then

$\phi(E(X)) \le E[\phi(X)]$.

Moment: $E(X^r)$ = r-th moment.

$E((X-\mu)^r)$ central moment.

mean. Variance. ~~Kurtosis.~~ skewness. Kurtosis.

 controlling. the tail heavy.

Multivariate moments. $E\begin{pmatrix} X \\ X \end{pmatrix} = \begin{pmatrix} EX_1 \\ EX_2 \end{pmatrix}$ $\operatorname{var}(X) = E\left((X - E(X))(X - E(X))'\right)$

$= E(XX') - E(X)E(X)'$

Conditional expectation $E(X_1 \mid X_2) = \int x_1 f(x_1 \mid x_2)\, dx_1$

~~$LIE: E(E(X_1 \mid X_2)) = E(X_1)$. martingale. mds~~

$E(X_{t+j} \mid X_{j-1}, X_{t-2}, \dots) = X_{j-1}$.

E(male | female) total score

LIE: $E(E(X_1 \mid X_2)) = E(X_1)$. $\int x_1 \int f(x_1, x_2)\, dx_2$

$= \int x_1 f(x_1)\, dx_1$

properties. pf of $\int\left[\int x_1 f(x_1 \mid x_2)\, dx_1\right] f(x_2)\, dx_2$

$= \frac{f(x_1, x_2)}{f(x_2)}$. $\int\int x_1 f(x_1)\, dx_1\, dx_2$

Convergent sequence -

1. Convergence in prob -

2. Convergence in m.s.

3. Convergence in distribution.

Slutsky Theorem.

~~delta method.~~ Lecture 2

~~Continuous mapping theorem,~~

* Chebyshev LLN.

* Lindeberg CLT.

Estimate: ~~(2) MLE. σ. μ poisson.~~ $\frac{e^{-\lambda}\lambda^k}{k!}$

(1) ~~x~~ ~~Co~~ method of ~~mo~~ moments.

$\frac{1}{n}\sum X_i$

$\frac{1}{n}\sum(X_i-\bar{X})^2$

evaluate: bias $E(\hat{\theta})-\theta_0$. $var.(\hat{\theta})$

MSE: $E((\hat{\theta}-\theta_0)^2)$ = bias^2 + var.

Efficiency. linear weight for mean

(2) MLE. normal. poisson.

Asy. normality: normal dist. non-normal. MM.

Conditional MLE. $y = x_i\beta + \varepsilon_i$.

~~logistic~~

~~m~~ $\beta = (X'X)^{-1}X'y_i$ matrix notation

~~$A(x) = \frac{1}{1+e^{-x}}$~~

Point estimation. Interval estimation. CLT. asy. normality

~~notati~~ Avar ~~lim~~ $\lim_{n\to\infty} Var(\sqrt{n}(\hat{\theta}-\theta_0))$

distributional theory..

method of moments.

1. mean. variance. bias $\frac{1}{n}\sum(x_i-\bar{x})^2 = \frac{1}{n}X(I-\frac{1}{n}\iota\iota')X$.

efficiency. smallest variance among ~~bias~~ unbiased.

evaluation: bias. $E(\hat{\theta})-\theta$. here $\hat{\theta}=\frac{1}{n}\sum\bar{x}$.

variance $E((\hat{\theta}_i-E(\hat{\theta}_i))^2$ $\theta_0=\mu$.

2. MLE. ~~KL~~ Kullback-Leibler distance.

$E_p(\log\frac{dP}{dQ})$ P.Q.

normal. poisson. $e^{-\lambda}\lambda^k/k!$

$\ell(\theta|x_1,\dots,x_n)=f(x_1,\dots,x_n|\theta)$

~~varian~~ sample variance $\sqrt{n}(S^2-\sigma^2)\xrightarrow{d}N(0,E[(X-\mu)^4]-\sigma^4)$.

if normal. $\sqrt{n}(\frac{S^2}{\sigma^2}-1)\xrightarrow{d}N(0,3-1)$.

Conditional MLE. $f(y|x_i;\beta)\sim N(x_i\beta,\Sigma^2)$.

3. Interval estimation:

normal distribution, $\frac{1}{n}\sum x_i$. finite sample distribution $N(0,-)$

$Pr(C_0\le\theta_0\le C_1)=0.95$. 1.96.

Hypothesis is a statement about the parameter space Θ

null is a subset. H. test is a decision to accept or not

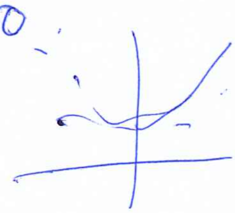
test function is a mapping ~~ϕ~~ $\phi: X^n \to \{0,1\}$.

power function $\beta_\phi(\theta) = E_\theta[\phi(X)]$. size. $\sup_{\theta \in \Theta} \beta_\phi(\theta)$.

~~CI $U(x)$ $L(x)$. s.t. $P_\theta(U(x) \leq \theta \leq L(x))$.~~

~~p-value~~ To Construct a test:

understand the distribution under the null hypo.



Consistent: if $\beta_\phi(\theta) \to 1$ for all $\theta \in \Theta_A$.

Eg. Trivial test

	accept	rej
null		
null false		

Joint linear hypotheses.

if $\sqrt{n}\begin{pmatrix}\hat{\theta}_1 - \theta_1 \\ \hat{\theta}_2 - \theta_2\end{pmatrix} \sim N(0, \Omega)$,

the $\theta_1 + \theta_2 = 0$?

$\sqrt{n} R \begin{pmatrix}\hat{\theta}_1 - \theta_1 \\ \hat{\theta}_2 - \theta_2\end{pmatrix} \sim N(0, R\Omega R')$

vector $\sqrt{n}(\hat{\theta} - \theta_0) \sim N(0, \Omega)$, then $\sqrt{n}R(\hat{\theta} - \theta) \sim N(0, R\Omega R')$

~~$n \cdot R(\hat{\theta} - \theta)$~~

$n \cdot R(\hat{\theta} - \theta)'(R\Omega R')^{-1} R(\hat{\theta} - \theta) \xrightarrow{d} \chi^2(\#\text{ restrictions})$

if $X \sim N(0, \Omega)$, then $X'\Omega^{-1}X \overset{d}{\sim} N(0, k)$