Policy - Data Analysis

Monday, July 17, 2017 10:55 AM

Topics
1. Data representation (Excel?)
2. Probability
3. Parameter Estimation – Point & interval
4. Hypothesis testing

Reference: An introduction to Probability & Statistics - V.Rohatgi & Saleh
Introduction to Probability & Statistics for Engineers & Scientists – Sheldon Ross

6 Quizzes – each 20%
↳ Best 5

80% Attendance - Strict

Problem Set 3

Data Analysis and Interpretation (EE 223)

Instructor: Prof. Prasanna Chaporkar

EE Department, IIT Bombay

1. The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} ce^{-x}e^{-2y} & 0 < x < \infty,\ 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = 1 \rightarrow$ calc. c

Compute (a) c, (b) $P\{X > 1, Y < 1\}$, (c) $P\{X < Y\}$, and (d) $P\{X < a\}$.

$\rightarrow 2$

2/3, but method

2. The joint density of X and Y is given by

$$f(x,y) = \begin{cases} e^{-(x+y)} & 0 < x < \infty,\ 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of the random variable X/Y.

Rigorous Treatment $\frac{25}{36} \rightarrow$

3. Sonia and Narendra decide to meet at a certain location. If each of them independently arrives at a time uniformly distributed between 5 PM and 6 PM, find the probability that the first to arrive has to wait longer than 10 minutes.

4. If the joint density function of X and Y is

$$f(x,y) = 6e^{-2x}e^{-3y} \quad 0 < x < \infty,\ 0 < y < \infty$$

and is equal to 0 outside this region, are the random variables independent? $\rightarrow$ yes What if the joint density function is

$$f(x,y) = 24xy \quad 0 < x < 1,\ 0 < y < 1,\ 0 < x + y < 1$$

and is equal to 0 otherwise? $\rightarrow$ No

5. Nitish and Lalu shoot at a target. The distance of each shot from the center of the target is uniformly distributed on $(0,1)$, independently of the other shot. What is the PDF of the distance of the losing shot from the center?

rigor? $\rightarrow$ 6. Let X, Y, Z be independent and uniformly distributed over $(0,1)$. Compute $P\{X \geq YZ\}$. (3/4)

7. Sum of two independent random variables

(a) If X and Y are independent random variables, both uniformly distributed on $(0,1)$, calculate the probability density of $X + Y$.

(b) If X and Y are independent Gamma random variables with respective parameters $(\alpha_1,\ \beta)$ and $(\alpha_2,\ \beta)$, then prove that $X + Y$ is also a Gamma random variable with parameters $(\alpha_1 + \alpha_2,\ \beta)$. Recall that a Gamma random variable has a density of the form

$$f(x;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \qquad x, \alpha, \beta > 0.$$

(c) If X and Y are independent Gaussian random variables with respective parameters $(\mu_1,\ \sigma_1^2)$ and $(\mu_2,\ \sigma_2^2)$, then prove that $X + Y$ is also a Gaussian random variable with parameters $(\mu_1 + \mu_2,\ \sigma_1^2 + \sigma_2^2)$. Recall that a Gaussian random variable has a density of the form

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2}}.$$

8. Conditional distribution of random variable

Q8 (a) $P(X>1 \mid Y=y) = 1 - P(X \leq 1 \mid Y=y)$

$= 1 - \lim_{\Delta y \to 0} \frac{P(X\leq 1,\ y<Y\leq y+\Delta y)/\Delta y}{P(y<Y\leq y+\Delta y)/\Delta y}$

$= 1 - \frac{\frac{\partial}{\partial y} F_{XY}(1,y)}{f_Y(y)}$

$= 1 + (-1 + e^{-1/y}) = e^{-1/y}$

$\frac{\partial}{\partial y}\int_0^y e^{-x}(1-e^{-1/x})\,dx = e^{-y}(1-e^{-1/y})$

$\frac{\partial}{\partial x}\int_a^x f(u)\,du = f(x)$

Q1 (b) $P\{X>1, Y<1\} = F_{XY}(X>1, Y\leq 1)$

$= \int_1^{\infty}\int_{-\infty}^{1} f_{XY}(x,y)\,dy\,dx = \left(\int_1^{\infty} 2e^{-x}\,dx\right)\left(\int_0^1 e^{-2y}\,dy\right)$

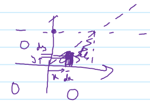
$= 2(e^{-1})\left(\frac{1}{2}\cdot(1-e^{-2})\right) = \frac{1}{e}\left(\frac{e^2-1}{e^2}\right) = \frac{e^2-1}{e^3}$

(d) $P(X<a) = \int_{-\infty}^{a}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dy\,dx = 0$ for $a \leq 0$

(for $a>0$) $= \int_0^a\int_0^{\infty} 2e^{-x}e^{-2y}\,dy\,dx = \int_0^a e^{-x}\,dx = e^{-x}\Big|_a^0 = (1-e^{-a}) = \frac{e^a-1}{e^a}$

$P(X<a) = \frac{e^a-1}{e^a}$

(c) $P(X<Y) = \int_0^{\infty}\int_0^{x} f_{XY}(x,y)\,dy\,dx$



$= \int_0^{\infty} e^{-x}(1-e^{-2x})\,dx$

$= \int_0^{\infty} e^{-x}\,dx - \frac{1}{3}\int_0^{\infty} e^{-3x}\,dx$

$= 1 - \frac{1}{3} = 2/3$

Q2 $z = X/y$ $\quad f_{XY}(x,y) = \begin{cases} e^{-(x+y)} & \text{in 1st quadrant} \\ 0 & \text{elsewhere (inc axes)} \end{cases}$

Let $w = x$; roots of this are $(w, \frac{w}{z})$; $J = \begin{bmatrix} (z/w) & (-z^2/w) \\ 1 & 0 \end{bmatrix}$

$f_{ZW}(z,w) = \frac{f_{XY}(w, \frac{w}{z})}{\begin{vmatrix} z/w & (-z^2/w) \\ 1 & 0 \end{vmatrix}} = \begin{cases} e^{-(w+w/z)}\frac{w}{z^2} & \text{for } 0<w<\infty \text{ and } 0<z<\infty \\ 0 & \text{elsewhere} \end{cases}$

$f_Z(z) = \int_{-\infty}^{\infty} f_{ZW}\,dw = \int_0^{\infty}\frac{1}{z^2} w e^{-w(1+1/z)}\,dw$; $\int x e^{-ax}\,dx = x\frac{e^{-ax}}{-a} - \frac{e^{-ax}}{a^2}$

$= \frac{1}{z^2}\left[0 + \left(0 + \frac{z^2}{(1+z)^2}\right)\right] = \frac{1}{(1+z)^2}$

Q3 S, N

$f_N(n), f_S(s) = \begin{cases} 1/60 \ ; & 0<n,s<60 \\ 0 \ ; & \text{otherwise} \end{cases}$; $f_{SN}(s,n) = \begin{cases} 1/3600 \ ; & 0<n,s<60 \\ 0 \ ; & \text{elsewhere} \end{cases}$

$P(|S-N| \leq 10) = P(S-N\leq 10 \text{ and } N-S \leq 10)$

$= P($

$f_Y(y) = \int_{-\infty}^{\infty} f_Y(y \mid X=x) f_X(x)\,dx = \int_0^{60} f_Y(y\mid X=x)\frac{1}{60}\,dx = \int_0^{10}\frac{(x+10)}{3600} + \int_{10}^{50}\frac{1}{36} + \int_{50}^{60}\frac{(70-x)}{3600}$

$\frac{1}{24} + \frac{2}{9} + \frac{1}{4} = \frac{3}{36} + \frac{8}{36} = \frac{11}{36}$

Q4 $f_{XY} = 6e^{-2x}e^{-3y}$ in 1st Quad

$f_X = \int_0^{\infty} 6e^{-2x}e^{-3y}\,dy = 2e^{-2x}\ e^{-3y}\Big|_{\infty}^{0} = 2e^{-2x}$

$f_{XY} = f_X \cdot f_Y \ \therefore X \perp\!\!\!\perp Y$

$f_{XY} = 24xy$ in ; $f_X(x) = \begin{cases} 0 \ ; & x \leq 0 \\ 12x(1-x)^2, & 0<x<1 \\ 0 \ ; & x \geq 1 \end{cases}$ $\quad \int_0^{1-x} 24xy\,dy = 12x\ y^2\Big|_0^{1-x} = 12x(1-x)^2$

(a) Suppose that the joint density of X and Y is given by

$$f(x,y) = \begin{cases} \dfrac{e^{-x/y}e^{-y}}{y} & 0 < x < \infty,\ 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find $P\{X > 1 | Y = y\}$. $e^{-1/y}$

(b) If X and Y are independent Poisson random variables with respective parameters λ_1 and λ_2. Prove that the conditional distribution of X given that $X + Y = n$ is a binomial distribution. A Poisson random variable has a pmf as

$$f(k;\lambda) = P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}.$$

(c) Rajan goes to the bank to make a deposit, and is equally likely to find 0 or 1 customer ahead of him. The times of service of these customers are independent and exponentially distributed with parameter λ. What is the CDF of Rajan's waiting time? Recall that a Exponential random variable has a density of the form

$$f(x;\lambda) = \lambda e^{-\lambda x} \quad x \geq 0.$$

9. Joint probability distribution of functions of random variables

(a) X and Y have joint density function

$$f(x,y) = \frac{1}{x^2y^2} \quad x \geq 1\ y \geq 1$$

Compute the joint density function of $U = XY$, $V = X/Y$. What are the marginal densities?

(b) Let X be exponentially distributed with mean 1. Once we observe the experimental value x of X, we generate a normal random variable Y with zero mean and variance $x + 1$. What is the joint PDF of X and Y?

$P(Y|X=x) =$

Q9 (a) $u = xy \quad v = x/y \quad x = vy$

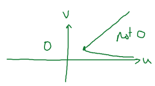
$vy^2 = u \Rightarrow y = +\sqrt{\frac{u}{v}}\ ;\ x = \sqrt{uv}$

$J = \begin{bmatrix} y & x \\ \frac{1}{y} & -\frac{x}{y^2} \end{bmatrix}$; $|J| = \left|-\frac{x}{y} - \frac{x}{y}\right| = \left|-\frac{2x}{y}\right| = +2\frac{\sqrt{uv}}{\sqrt{\frac{u}{v}}} = +2v$

$f_{UV}(u,v) = \dfrac{\frac{1}{(u\cancel{v})}\frac{\cancel{v}}{u}}{+2v} = \dfrac{+1}{2u^2v}$; $u \geq 1$; $\frac{1}{u} \leq v \leq u$

$f_U(u) = \int_{\frac{1}{u}}^{u} \frac{+1}{2u^2v}\,dv = \frac{+1}{2u^2} \ln v\Big|_{\frac{1}{u}}^{u} = \frac{+1}{u^2}\,(2\ln u = +\frac{\cancel{2}\ln u}{\cancel{2}u^2}$

$f_U(u) = \begin{cases} \frac{\ln u}{u^2} & \text{for } u \geq 1 \\ 0 & \text{for } u < 1 \end{cases}$



$f_V(u,v) = \begin{cases} 0 & \text{for } v \leq 0 \\ \frac{1}{2} & \text{for } 0 < v < 1 \\ \frac{1}{2v^2} & \text{for } 1 \leq v \end{cases}$

$\int_{\frac{1}{v}}^{\infty} \frac{1}{2u^2v}\,du = \frac{1}{2v}\left(\frac{+1}{u}\right)\Big|_{\infty}^{1/v} = \frac{1}{2v}(v - 0) = \frac{1}{2}$

$\int_{v}^{\infty} \frac{1}{2u^2v}\,dv = \frac{1}{2v}\ \frac{1}{u}\Big|_{\infty}^{v} = \frac{1}{2v}\left(\frac{1}{v}\right) = \frac{1}{2v^2}$

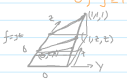
Q9 (b) $f_X(x) = \begin{cases} e^{-x} ; & x \geq 0 \\ 0 , & x < 0 \end{cases}$ | $f_Y(y|X=x) = \frac{1}{f_X(x)} \int_{-\infty}^{y} f_{XY}(x,y')\,dy'$

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

$f_{XY} = f_X(x) \cdot f_{Y|X}(y|x) = e^{-x} \frac{1}{\sqrt{2\pi}\,(x+1)} \exp\left(\frac{-y^2}{2(x+1)^2}\right)$

$f_Y(y) = \begin{cases} 0 ; & y \leq 0 \\ 12y(1-y)^2 ; & 0 < y < 1 \\ 0 ; & y \geq 1 \end{cases}$ $\therefore$ Not indep.

Q6



$P(X > YZ) = \frac{V_{above}}{V_T} = 1 - \frac{V_{below}}{V_T}$

$= \iint (yz)\,dy\,dz$

$\left(\int_0^1 y\,dy\right)^2 = \frac{1}{4}$; $V_T = 1$

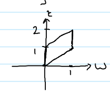
$\therefore P(X > YZ) = 3/4$

Q7 (a) $f_X = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{else} \end{cases}$ $f_Y = \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{else} \end{cases}$ – indep

$f_{XY} = \begin{cases} 1 , & 0 < x,y < 1 \\ 0 , & \text{else} \end{cases}$

$z = X + Y \quad w = X$; $J = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}$; $(w, z-w)$ is the only root

$f_{WZ} = -\begin{cases} 1 , & 0 < w < 1 \text{ and } 0 < z - w < 1 \\ 0 ; & \text{else} \end{cases}$



$f_Z = \int f_{WZ}\,dw = \begin{cases} 0 , & z \leq 0 \\ z , & 0 < z \leq 1 \\ 2 - z , & 1 < z < 2 \\ 0 , & 2 \leq z \end{cases}$

Q8 (c) Rajan's wait is a r.v. X ; # of people is an indicator r.v., say Y

$F_X(x|Y=1) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geq 0 , \\ 0 & \text{for } x < 0 \end{cases}$

$F_X(x|Y=0) = \begin{cases} 1 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$

$F_X(x) = F_X(x|Y=0) \cdot P(Y=0) + F_X(x|Y=1)\ P(Y=1)$

$= \begin{cases} 0 & \text{for } x < 0 \\ 1 - \frac{1}{2}e^{-\lambda x} & \text{for } x \geq 0 \end{cases}$

Q8 (b) $P(X=k) = e^{-\lambda_1}\frac{\lambda_1^k}{k!}$; $P(Y=k) = e^{-\lambda_2}\frac{\lambda_2^k}{k!}$; $X \perp\!\!\!\perp Y$

$k = 0,1,2,\ldots$

$P(X | X+Y=n) = \dfrac{P(X=k, Y=n-k)}{P(X+Y=n)}$

$= \dfrac{\cancel{e^{-(\lambda_1+\lambda_2)}}\frac{\lambda_1^k}{k!} \times \frac{\lambda_2^{n-k}}{(n-k)!}}{\frac{\cancel{e^{-(\lambda_1+\lambda_2)}}}{n!}(\lambda_1+\lambda_2)^n}$

$= \dfrac{{}^nC_k\ \lambda_1^k\ \lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n}$, which is Binomial

$P(X+Y=n) = \sum_{k=0}^{n} P(X=n-k | Y=k) \cdot P(Y=k)$

$= \sum_{k=0}^{n}\left(e^{-\lambda_2}\frac{\lambda_2^k}{k!}\right) \times \left(e^{-\lambda_1}\frac{\lambda_1^{n-k}}{(n-k)!}\right) = \frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{(n-k)!\,k!}\lambda_1^k\lambda_2^{n-k}$

$= \frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n$

Q5 X & Y ,

$F_X, F_Y = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & 1 < x \end{cases}$ | $f_X, f_Y = \begin{cases} 0 & x < 0 \ \& \ 1 < x \\ 1 & 0 \leq x \leq 1 \end{cases}$

$X \perp\!\!\!\perp Y$

$F_{XY} = \begin{cases} 0 & x < 0 \text{ or } y < 0 \\ xy & 0 \leq x,y \leq 1 \\ y & x > 1 \ \& \ 0 \leq y \leq 1 \\ x & y > 1 \ \& \ 0 \leq x \leq 1 \\ 1 & x > 1 \ \& \ y > 1 \end{cases}$

$Z = \max\{X, Y\}$

$P(X \leq z) = P(Y \leq z) = z$

$F_Z(z) = P(\max\{X,Y\} \leq z) = P(X \leq z, Y \leq z)$

$= P(X \leq z) \cdot P(Y \leq z) = z \cdot z = z^2$

$\therefore f_Z(z) = \begin{cases} 2z ; & 0 \leq z \leq 1 \\ 0 ; & z < 0 \text{ or } z > 1 \end{cases}$

Problem Set 4

Saturday, October 7, 2017 1:30 AM

Problem Set 2
Data Analysis and Interpretation (EE 223)

Instructor: Prof. Prasanna Chaporkar
EE Department, IIT Bombay

1. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with probability density function $f_{\vec{\theta}}(\cdot)$. Find the sufficient statistic in the following cases:
   (a) $f_\theta(x) = e^{\theta - x}$ for $x > \theta$ and 0 otherwise. $T(x_1, .., x_n) = \min\{x_n\}$
   (b) $f_\theta(x) = 1/(2\theta)$ for $x \in [1-\theta, 1+\theta]$ and 0 otherwise. $T(x_1, \ldots, x_n) = (\min\{x_k\}, \max\{x_k\})$
   (c) $f_\theta(x) = 1/\theta$ for $x \in [0, \theta]$ and 0 otherwise. $T = \max\{x_k\}$
   (d) $f_\theta(x) = c(\theta) x^{-\theta}$ for $x \geq 1$ and 0 otherwise. Here, $\theta \geq 2$ and $c(\theta)$ is a normalization constant depending on θ. $T = \prod_k x_k$

2. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with probability density function $f_{\vec{\theta}}(\cdot)$. Define,
(b)
$$k(\vec{x}, \vec{y}, \vec{\theta}) = \frac{f_{\vec{\theta}}(\vec{x})}{f_{\vec{\theta}}(\vec{y})}.$$
A sufficient statistic T is called minimal if the following holds: $k(\vec{x}, \vec{y}, \vec{\theta})$ does not depend on $\vec{\theta}$ if and only if $T(\vec{x}) = T(\vec{y})$. Show that
   (a) if $X_1 \sim G(\mu, 1)$, then $T(\vec{x}) = \sum_{k=1}^n x_k$ is minimal sufficient. Also, $\tilde{T}(\vec{x}) = (\sum_{k=1}^n x_k, \sum_{k=1}^n x_k^2)$ is not minimal. $\rightarrow k \neq fn(\theta)$ & $T(\bar{x}) \neq T(\bar{y})$ when $\Sigma x_i = \Sigma y_i$ but $\Sigma x_i^2 \neq \Sigma y_i^2$
   (b) Verify that all the sufficient statistics found in Question 1 are minimal.

3. Let $X \sim \text{Poisson}(\lambda)$, the parameter $\lambda \in (0, \infty)$, and $\psi(\lambda) = 1/\lambda$. Find an unbiased estimator for ψ.

4. You need to aid a physicist in estimating the rate at which a radio active material emits gamma particles. It is known that the interval between the two consecutive emissions is an exponential random variable with parameter λ. Moreover, inter-emission periods are independent. You choose to put the radio active material with a photographic plate in a lead container for T time units. At the end of this period, you take out the plate and measure the number of marks on the plate.
   (a) Show that only noting the number of marks on the photographic plate is sufficient to estimate the rate
   (b) Give an unbiased estimator for λ $T = \Sigma x_i$ is sufficient

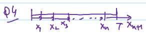
5. You are tasked with approximating the number of tigers in a tiger reserve. You install sensors near a water body that can uniquely identify a tiger that comes close to the water body. Information from the locals allows you to believe that each tiger visits the water body with probability 0.1 independent of other tigers.
   (a) Show that counting the number of unique tigers that visited the water body is sufficient to estimate the number of tigers
   (b) Provide an unbiased estimator

6. Consider random variables X and Y with joint probability density function $f_{XY}(x, y)$, and marginals $f_X(x)$ and $f_Y(y)$. Let $Z = E[X|Y]$ be a random variable from $(\Re, \mathcal{B}(\Re))$ to $(\Re, \mathcal{B}(\Re))$ defined as follows: for $y \in \Re$, $Z(y) = E[X|Y = y]$. Argue that
   (a)
$$Z = \frac{1}{f_Y(Y)} \int_{-\infty}^{\infty} x f_{XY}(x, Y) dx$$
   (b) Find $E[Z]$. $= E[X]$

Q3/ $E_\theta[\delta(\bar{x})] = \psi(\theta) \Rightarrow \sum_{i=0}^{\infty} \delta(i) \frac{\lambda^i}{\lfloor i} e^{-\lambda} = \frac{1}{\lambda}$

$\Rightarrow \sum_{i=0}^{\infty} \delta(i) \frac{\lambda^{i+1}}{\lfloor i} = e^{\lambda} \Rightarrow \left(\frac{\delta(0)\lambda}{1} + \frac{\delta(1)\lambda^2}{1} + \frac{\delta(2)\lambda^3}{\lfloor 2} + \ldots \right) = \left(1 + \frac{\lambda}{1} + \frac{\lambda^2}{\lfloor 2} + \frac{\lambda^3}{\lfloor 3} + \ldots \right)$

Constant term in RHS but not LHS, So They cannot be equal

Q4 

$Y = \sum_{i=1}^{k} X_i$ $Y \sim \text{Gamma} \rightarrow \lambda^k e^{-\lambda x} \frac{x^{k-1}}{\lfloor k-1}$; $k = \#$ of Marks on plate

$P_\lambda(N=k) = P(\sum_{i=1}^{k} X_i \leq T, \sum_{i=1}^{k+1} X_i > T)$

$= P(Y \leq T, Y+X > T) = \int_{-\infty}^{\infty} P(Y \leq T, Y+X > T / Y=y) f_Y(y) dy$

$= \int_0^T P(X > T-y) f_Y(y) dy = \int_0^T \exp(\lambda(y-T)) \lambda^k e^{-\lambda y} \frac{y^{k-1}}{\lfloor k-1} dy = \frac{\lambda^k}{\lfloor k-1} \int_0^T \exp(-\lambda T) y^{k-1} dy = \frac{\lambda^k}{\lfloor k-1} e^{-\lambda T} \frac{y^k}{k} \Big|_0^T = \frac{\lambda^k e^{-\lambda T} T^k}{\lfloor k}$

$P_\lambda(N=k) = e^{-\lambda T} \left(\frac{\lambda^k T^k}{\lfloor k} \right)$; $h(\bar{x}) = \frac{T^k}{\lfloor k}$, $g_\lambda(k) = \lambda^k e^{-\lambda T}$ $\therefore$ knowing 'k' is sufficient

$\sum_i \delta(k) P(x_i = k) = \lambda$; $\psi(\lambda) = \lambda$; $\delta_1(\bar{x}) =$ $E[\delta_1] =$

Q5 (a) $p = 0.1$, r visits, i.e. $\Sigma x_i = r$; $n = \#$ of tigers in forest

$P(\Sigma x_i = r) = {}^nC_r \, p^r (1-p)^{n-r}$

$h(\bar{x}) = p^r$ $g_n(r) = {}^nC_r (1-p)^{n-r}$

(b) $\psi(n) = n$; $\delta_1(\bar{x}) = \frac{\Sigma x_i}{p}$

Q6 $Z(y) = E[X | Y=y] = \int_{-\infty}^{\infty} x \, f_{X|Y}(x, y) = \int_{-\infty}^{\infty} x \frac{f_{XY}(x,y)}{f_Y(y)} dx = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} x f_{XY}(x,y) dx$

(b) $E[Z] = \int_{-\infty}^{\infty} \left(\frac{1}{f(y)} \int_{-\infty}^{\infty} x f_{XY}(x,y) dx \right) f_Y(y) dy = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{XY}(x,y) dy \, dx = \int_{-\infty}^{\infty} x f_X(x) dx = E[X]$

7. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with probability density function $f_{\vec{\theta}}(\cdot)$. Define $X = X_1$ and $Y = \sum_{k=1}^{n} X_k$.

   (a) For $n = 3$ and let the density function be exponential with parameter $\lambda > 0$. Find $E[X|Y]$. Do explicit calculations.

   (b) Find $E[X|Y]$ for the general density function.

8. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with probability density function $f_{\vec{\theta}}(\cdot)$. Let T be a sufficient statistic and let δ be any unbiased estimate of the given function ψ. Show that $E_\theta[\delta|T]$ is also an unbiased estimate of ψ.

9. Show that the sample variance is an unbiased estimator for the variance. $\psi = \sigma^2$; $\delta = \sum \frac{(x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$; $E(\delta) = \frac{\sum E[x_i^2] - E[n\bar{x}^2]}{n-1} = \frac{(\sigma^2+\mu^2)n - (\sigma^2 + n\mu^2)}{n-1} = \sigma^2$; $\operatorname{var}(X) = E[X^2] - (E[X])^2 \Rightarrow E[X^2] = \sigma^2 + \mu^2$

   $\operatorname{Var}(\bar{X}) = E[\bar{X}^2] - (E[\bar{X}])^2 \Rightarrow E[\bar{X}^2] = \mu^2 + \frac{\sigma^2}{n}$

10. Find unbiased estimate for $\sigma > 0$, where samples are drawn from $G(\sigma, \sigma)$. $f_X = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(x-\sigma)^2}{2\sigma^2}\right)$

$x_1, \ldots, x_n \quad \psi(\sigma) = \sigma$

$f_{\bar{X}}(\bar{x}) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma}\right)^n \exp\left(\frac{\sum x_i^2}{2\sigma^2} - \frac{1}{\sigma}\sum x_i + \frac{1}{2}\right)$

$\delta(\bar{x}) = \frac{\sum x_i}{n}$

$E[\delta(\bar{x})] = \frac{n\sigma}{n} = \sigma$

Q8 $E[E_\theta(\delta|T)] = E(\delta) \rightarrow$ To prove

$E[E(X|Y)] = E\left[\sum_x x P(X=x|Y)\right] = \sum_y \left(\sum_x x P(X=x|Y)\right) P(Y=y)$

$= \sum_y \sum_x x P(X=x, Y=y) = \sum_x x \sum_y P(X=x, Y=y)$

$= \sum_x x P(X) = E[X]$ — (1) for discrete, from Q6 for cont.

$E[\delta|T]$ is independent of θ — (2)

from (1) & (2), $E[E[\delta|T]] = E[\delta]$

Q7 $\sum_{i=1}^{3} E[X_i|Y] = E[Y|Y] = y$

$\Rightarrow E[X_1|Y] = y/3$

(a) $E[X_1|Y] = \int_{-\infty}^{\infty} x \frac{f_{X_1,Y}(x_1,y)}{f_Y(y)}\, dx$; $Y \sim \text{Gamma}(\lambda)$; $P(\ =y) = \lambda^k e^{-\lambda y} \frac{y^{k-1}}{\lfloor k-1}$

$f_X \quad \frac{f_{X_1,Y}(x_1,y)}{f_X} = f_X(x) \; f_{Y|X_1}(Y|X_1 = x_1)$

Problem Set 5

Sunday, October 29, 2017 11:09 AM

Problem Set 5
Data Analysis and Interpretation (EE 223)

Instructor: Prof. Prasanna Chaporkar
EE Department, IIT Bombay

1. Let $X_1, \ldots, X_n$ be i.i.d. Poisson(λ). Find UMVUE for λ.
2. Let $X_1, \ldots, X_n$ be i.i.d. Poisson(λ).
   (a) Find $E_\lambda[X_1^2]$.
   (b) Find $E_\lambda[X_1^2 | \sum_{i=1}^n X_i = y]$.
   (c) Find $\psi(\lambda)$ s.t. $E_\lambda[X_1^2 | \sum_{i=1}^n X_i]$ is UMVUE for $\psi(\lambda)$.
3. Let $X_1, \ldots, X_n$ be i.i.d. Gaussian(μ, σ^2), where μ is known. Consider the following family of estimators for σ^2,
$$\delta_K(X_1, \ldots, X_n) = \frac{1}{K}\sum_{i=1}^n (X_i - \bar{X})^2,$$
   where $\bar{X}$ is the sample mean.
   (a) Find MSE for $\delta_K(\cdot)$.
   (b) Find the optimal value of K for which MSE is the minimum.
4. Let $X_1, \ldots, X_n$ be i.i.d. RVs with $f_\lambda(\cdot)$ s.t.
$$f_\lambda(x) = \frac{1}{\lambda}e^{-\frac{x}{\lambda}}, \quad \forall\, x \geq 0,\ \lambda \in (0, \infty)$$
   Let $\psi(\lambda) = \lambda^2$. Find UMVUE for $\psi(\cdot)$.
5. Let $X_1, \ldots, X_n$ be i.i.d. RVs with $f_\theta(\cdot)$ where
$$\begin{aligned} f_\theta(x) &= 2x/\theta^2, \quad 0 < x < \theta, \\ &= 0, \quad \text{otherwise.} \end{aligned}$$
   Let $\delta_c(\bar{x}) = c\max\{x_1, \ldots, x_n\}$.
   (a) Find MSE for $\delta_c(\bar{x})$.
   (b) Find c that minimizes MSE.
6. Let $X \sim N(\theta, \theta^2)$ and $\theta \in [0, \infty)$. Find MLE for θ^2.
7. Let $X_1, \ldots, X_n$ be i.i.d. RVs. Find MLE for θ.
   (a) Bernoulli distribution with parameter θ
   (b) Geometric distribution
$$f_\theta(x) = (1-\theta)^{x-1}\theta$$
   (c) Poisson distribution
$$f_\theta(x) = \frac{\theta^x e^{-\theta}}{x!}$$
   (d) Binomial distribution
$$f_\theta(x) = \frac{n!}{x!(n-x)!}\theta^x(1-\theta)^{n-x}$$

Q1 $f_{X_1}(x_1) = \dfrac{e^{-\lambda}\lambda^{x_1}}{x_1!}$; $\psi = \lambda$

$E[X_1] = \lambda$; $T = \frac{1}{n}\sum_i x_i$; $E[T] = n \times \frac{1}{n} \times \lambda = \lambda$

T is unbiased, If T is complete, then it is UMVUE

$E(g(T)) = 0 \Rightarrow \sum_k g(T) P(\sum x_i = nk) = 0 \Rightarrow \sum_k g(T) \underbrace{e^{-n\lambda}\frac{(n\lambda)^{nk}}{nk!}}_{>0\ \forall k} = 0$ $\therefore g(x) = 0\ \forall x$

$\therefore$ T is complete, hence it is UMVUE

Q2 (a) $E[X_1^2] = \sum_k e^{-\lambda}\dfrac{\lambda^{k-2}}{(k-2)!}\times\lambda^2 + e^{-\lambda}\dfrac{\lambda^{k-1}}{(k-1)!}\lambda$

$= \lambda^2 + \lambda$

(b) $E[X_1^2 | \sum_{i=1}^n X_i = y] = \sum_{k=0}^{y} k^2 P(X_1 = k | \sum X_i = y) = \sum_0^y k^2 \dfrac{P(X_1 = k \;\&\; \sum_2^n X_i = y-k)}{P(\sum X_i = y)} = \sum_1^y \dfrac{(n-1)^{y-k}}{n^y}\,{}^{y-1}C_{k-1}\, y + \sum_2^y \dfrac{(n-1)^{y-k}}{n^y}\,{}^{y-2}C_{k-2}\, y(y-1)$

$+ \left(\frac{n-1}{n}\right)^y y + \left(\frac{n-1}{n}\right)^y \frac{1}{(n-1)} y(y-1)$

$= \dfrac{y}{n^2}(n + (y-1)) = \dfrac{y(y-1)+ny}{n^2}$

(c) $\psi(\lambda) = E_\lambda(X_1^2) = \lambda^2 + \lambda$

Q4 $f_\lambda(x) = \begin{cases} \frac{1}{\lambda}e^{-x/\lambda} ; & x \geq 0 ; \lambda > 0 \\ 0, & \text{else} \end{cases}$; $\psi(\lambda) = \lambda^2$ | $E(X_1) = \int_0^\infty \lambda\, t e^{-t} dt = \lambda\left[-te^{-t} - e^{-t}\right]_0^\infty = \lambda[0 - (0-1)] = \lambda$

$E\left[\frac{X_1^2}{2}\right] = \int_0^\infty \frac{1}{2}\lambda^2 t^2 e^{-t}\, dt = \frac{\lambda^2}{2}\left[-t^2e^{-t} - 2te^{-t} - 2e^{-t}\right]_0^\infty = \frac{\lambda^2}{2}[0 - (-2)] = \lambda^2$

$f_\lambda(\bar{x}) = \left(\frac{1}{\lambda}\right)^n 1_{\min\{x_n\}\geq 0}\, e^{-\frac{1}{\lambda}(\sum x_i)}$; $T = \sum X_i$ is sufficient, to show complete,

$\int_0^\infty g(T)\frac{1}{\lambda^n}e^{-\sum x/\lambda}dx = 0 \Rightarrow g(T) = 0$: T is complete

$\delta(\bar{x}) = \frac{1}{2n}\sum X_i^2$; $E(\delta) = \frac{1}{2n}2\lambda^2 n = \lambda^2 = \psi$; $\delta(\bar{x})$ is unbiased

$\therefore E[\delta | T]$ is UMVUE

Q5 $f_\theta(\bar{x}) = 1_{\min\{x_i\}\geq 0}\, 1_{\max\{x_i\}\leq\theta}\, \dfrac{2^n \prod x_i}{\theta^{2n}}$; $T(\bar{x}) = (\max\{x_n\}, \prod x_i)$

$E[X_1] = \int_0^\theta x \frac{2x}{\theta^2}dx = \frac{2}{\theta^2}\frac{x^3}{3}\Big|_0^\theta = \frac{2}{3}\theta$; $Y = \max\{x_i\}$

$\delta_c = c\max\{x_i\}$; $MSE = E[(\delta_c - \psi)^2] = E[\delta^2 + \psi^2 - 2\delta\psi]$

$P_\theta(Y \leq y) = P_\theta(X_i < Y) = \begin{cases} 1 ; & \text{if } y \geq \theta \\ 0 ; & \text{if } y < 0 \\ \frac{y^{2n}}{\theta^{2n}} ; & \text{else} \end{cases}$ | $f_Y = \begin{cases} 2n\, y^{2n-1}/\theta^{2n} ; & 0 \leq y \leq \theta \\ 0 ; & \text{else} \end{cases}$

$E[\delta] = c\frac{2n}{\theta^{2n}}\int_0^\theta y^{2n} dy = \frac{2nc\,\theta}{2n+1}$

δ_c is unbiased for $c = \frac{2n+1}{2n}$, which minimises MSE

Q6 $X \sim N[\theta, \theta^2]$, $\theta \in [0, \infty)$

$F = \dfrac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$; $L = (2\pi)^{-n/2}\theta^{-n}\exp\left(\dfrac{-\sum(x-\theta)^2}{2\theta^2}\right)$

$\ln L = (\ ) - n\ln\theta - \dfrac{\sum x^2}{2\theta^2} - \dfrac{n}{2} + \dfrac{1}{\theta}\sum x_i$

$0 = 0 - n\dfrac{1}{\theta} + \dfrac{\sum x^2}{\theta^3} + 0 - \dfrac{\sum x_i}{\theta^2}$

$n\theta^2 + (\sum x_i)\theta - (\sum x_i^2) = 0$

(e) Negative Binomial distribution

$$f_\theta(x) = \binom{x-1}{r-1} \cdot \theta^r \cdot (1-\theta)^{x-r}$$

(f) Exponential distribution

$$f_\theta(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \qquad x \geq 0$$

(g) Gaussian distribution

$$f_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\theta)^2/2\sigma^2}$$

(h) Rayleigh distribution

$$f_\theta(x) = \frac{x}{\theta^2} e^{\left(\frac{-x^2}{2\theta^2}\right)}, \qquad x > 0$$

(i) Gamma distribution

$$f_\theta(x) = \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x}, \qquad x > 0, \quad \theta, \alpha > 0.$$

(j) Pareto distribution

$$f_\theta(x) = \theta \frac{\beta^\theta}{x^{\theta+1}}, \qquad x \geq \beta, \quad \theta, \beta > 0.$$

8. Let $X_1, \ldots, X_n$ be i.i.d. Poisson(λ). Find MLE for $(1-\lambda)^2$.

of the two roots, the one w/ $\frac{\partial^2 L}{\partial \theta^2} < 0$ is the answer

Q3 $\delta_k = \frac{1}{k} \sum_i (X_i - \bar{X})^2 = \frac{\sigma^2}{k} \chi^2_{n-1}$; $\psi(\sigma) = \sigma^2$

$$MSE = \frac{\sigma^4}{k^2} 2(n-1) + \left(\frac{\sigma^2}{k}(n-1) - \sigma^2\right)^2$$

to min. MSE, $\sigma^2 \cancel{\sigma^4} \cancel{2} \cancel{(n-1)} \left(\frac{2}{k}\right) + \cancel{2}\left(\quad \right)\left(\cancel{\sigma^4}\cancel{(n-1)}\right) = 0$

$$\frac{2\cancel{\sigma^2}}{k} + \frac{\cancel{\sigma^2}}{k}(n-1) - \cancel{\sigma^2} = 0$$

$$\frac{2+n-1}{k} = 1 \Rightarrow k = n+1$$

Problem Set 6

Thursday, November 9, 2017 9:49 AM

Problem Set 6
Data Analysis and Interpretation (EE 223)
Instructor: Prof. Prasanna Chaporkar
EE Department, IIT Bombay

1. Suppose $x_1, x_2, ..., x_n$ is a random sample from an exponential distribution with parameter θ. Is the hypothesis $H : \theta = 3$ a simple or a composite hypothesis? What about $H : \theta > 2$?
Composite

2. By using central limit theorem to approximate the distribution of $\sum_{i=1}^{n} X_i$, show that the smallest value for n required to make $\alpha = 0.05$ and $\beta \le 0.1$ is approximately 213. Let α denote the Type I error probability and β denote the Type II error probability.
We need E_X & $var(X)$

3. Suppose X is a single observation from a population with probability density given by $f(x) = \theta x^{\theta-1}$ for $0 < x < 1$. Find the test with best critical region. That is, find the most powerful test, with significance level $\alpha = 0.05$, for testing a single null hypothesis $H_0 : \theta = 3$ against the simple alternative hypothesis $H_A : \theta = 2$.

4. Suppose $X_1, X_2, ..X_n$ is a random sample from a normal population with mean μ and variance 16. Find the test with best critical region, with a sample size of $n = 16$ and a significance level $\alpha = 0.05$ to test the null hypothesis $H_0 : \mu = 10$ against the alternative hypothesis $H_A : \mu = 15$.

5. $X = (X_1, X_2, ..., X_n)$ is a sequence of Bernoulli trials with unknown success probability θ, the likelihood
$$L(\theta|x) = (1-\theta)^n \left(\frac{\theta}{1-\theta}\right)^{x_1+x_2+...+x_n}$$
. For the test $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$, take $\theta_0 = 1/2, \theta > 1/2$ and $\alpha = 0.05$.

6. Suppose that $Y_1, Y_2, ..., Y_n$ are independent Poisson (λ) random variables and consider testing $H_0 : \lambda = \lambda_0$ against $H_1 : \lambda = \lambda_1$, where $\lambda_1 > \lambda_0$. Significance level $= \alpha$. Determine the decision regions using Neyman-Pearson lemma.

7. Let X be a random variable whose pmf under H_0 and H_1 is given by

x	1	2	3	4	5	6	7
$f(x/H_0)$	0.01	0.01	0.01	0.01	0.01	0.01	0.94
$f(x/H_1)$	0.06	0.05	0.04	0.03	0.02	0.01	0.79

Use Neyman-Pearson Lemma to find the most powerful test for H_0 versus H_1 with size $\alpha = 0.04$. Compute probability of Type II Error for this test.

8. The R.V X has the pdf $f(x) = e^{-x}, x > 0$. One observation is obtained on the R.V $Y = X^{\theta}$, and a test of $H_0 : \theta = 1$ versus $H_1 : \theta = 2$ needs to be constructed. Find the UMP level $\alpha = 0.1$ test and compute the Type II Error probability.

9. Show that for a random sample $X_1, X_2, ..., X_n$ from a $N(0, \sigma^2)$ population, the most powerful test of $H_0 : \sigma = \sigma_0$ versus $H_1 : \sigma = \sigma_1$, where $\sigma_0 < \sigma_1$ is given by
$$\phi(\Sigma x_i^2) = \begin{cases} 1, & \Sigma x_i^2 > c \\ 0, & x_i^2 \le c \end{cases}$$
For a given value of α, the size of Type I error, show how the value of c is explicitly determined.

Q3 $f_\theta(x) = \begin{cases} \theta x^{\theta-1} & ; 0<x<1 \\ 0 & ; \text{else} \end{cases}$

$H_0 : \theta = 3$; $H_1 : \theta = 2$

$\lambda = \frac{f_{\theta_1}}{f_{\theta_0}} = \frac{2x}{3x^2} = \frac{2}{3x}$

$\phi = \begin{cases} 1 ; & \text{if } x_i < k \\ \gamma ; & \text{if } x_i = k \\ 0 ; & \text{if } x_i > k \end{cases}$; $E_{\theta_0}[\phi] = 1 \times P(x_i < k) = \int_0^k 3x^2\,dx = k^3 \le \alpha$
$\Rightarrow \underline{k = \alpha^{1/3}}$

$P(x_i = k) = 0$, so γ is irrelevant

Q4 $f_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\theta)^2}{2\sigma}\right)$

$H_0 : \mu = 10$; $H_1 : \mu = 15$

$\frac{f_1(\bar{x})}{f_0(\bar{x})} = \exp\left(\sum_i \frac{(x_i-10)^2 - (x_i-15)^2}{2\sigma}\right) = \exp\left(\frac{1}{2\sigma}\sum_i (2x_i-25)(5)\right) = (\ )\exp\left(\frac{5}{\sigma}\sum_i x_i\right)$; inc fn. of Σx_i

$\phi = \begin{cases} 1 , & \Sigma x_i > k \\ \gamma ; & \Sigma x_i = k \\ 0 ; & \Sigma x_i < k \end{cases}$; $E_{\theta_0}[\phi(\bar{x})] = P_{\theta_0}(\Sigma x_i > k)$; $\frac{\Sigma x_i}{n} \sim G(\mu_0, \sigma)$

$Y = \frac{1}{\sqrt{\sigma}}\left(\frac{\Sigma x_i}{n} - \mu_0\right) > \left(\frac{k}{n} - \mu_0\right)\frac{1}{\sqrt{\sigma}}$; $Y \sim G(0,1)$

$= EFRC\left(\left(\frac{k}{n} - \mu_0\right)\frac{1}{\sqrt{\sigma}}\right) \le \alpha$

$\therefore \alpha = EFRC\left(\frac{1}{\sqrt{\sigma}}\left(\frac{k}{n} - \mu_0\right)\right) \to$ get k

γ is irrelevant

Q5 $L(\theta, \bar{x}) = (1-\theta)^n \left(\frac{\theta}{1-\theta}\right)^{\Sigma x_i}$

$\frac{f_{\theta_1}}{f_{\theta_0}} = \left(\frac{1-\theta_1}{1-\theta_0}\right)^n \left(\frac{\theta_1}{1-\theta_1} \times \frac{1-\theta_0}{\theta_0}\right)^{\Sigma x_i} = 2^n (1-\theta_1)^n \left(\frac{\theta_1}{1-\theta_1}\right)^{\Sigma x_i}$

$\theta_1 > \frac{1}{2} \Rightarrow \theta_1 > 1-\theta_1 \Rightarrow \frac{f_{\theta_1}}{f_{\theta_0}}(\bar{x})$ inc. w/ Σx_i

$\phi(\bar{x}) = \begin{cases} 1 ; & \Sigma x_i > k \\ \gamma ; & \Sigma x_i = k \\ 0 ; & \Sigma x_i < k \end{cases}$; $E_{\theta_0}[\phi(\bar{x})] = P_{\theta_0}(\Sigma x_i > k) + \gamma P_{\theta_0}(\Sigma x_i = k)$

$= \sum_{r \ge k+1}^{n} {}^nC_r \theta_0^r (1-\theta_0)^{n-r} + \gamma\, {}^nC_k \theta_0^k (1-\theta_0)^{n-k}$

$= \left(\frac{1}{2}\right)^n \left(\sum_{r=k+1}^{n} {}^nC_r + \gamma\, {}^nC_k\right) \le \alpha$

where k is st $\sum_{r=k+1}^{n} {}^nC_r \le 2^n \alpha < \sum_{r=k}^{n} {}^nC_r$

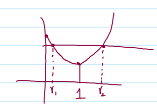
Q8 $f_{\theta_1}(y) = \frac{e^{-\sqrt{y}}}{2\sqrt{y}}$; $f_{\theta_0}(y) = e^{-y}$

$\lambda(\bar{y}) = \frac{e^{y-\sqrt{y}}}{2\sqrt{y}} \Rightarrow \ln\lambda = y - \sqrt{y} - \ln 2 - \frac{1}{2}\ln y$

$1 - \frac{1}{2\sqrt{y}} - 0 - \frac{1}{2}\frac{1}{y} = \frac{2y}{2y} - \frac{1}{2y} - \frac{\sqrt{y}}{2y} > 0 \Rightarrow \frac{2x^2 - x - 1}{2y} > 0$

$\frac{(x-1)(2x+1)}{y} > 0$

$\lambda'(y) \begin{cases} < 0 & \text{for } 0<y<1 \\ > 0 & \text{for } y>1 \end{cases}$



r_1 & r_2 are roots of $\lambda(\bar{y}) = k$

$\phi(\bar{y}) = \begin{cases} 1 ; & \text{if } \lambda(\bar{y}) > k \\ \gamma ; & \\ 0 ; & \text{if } \lambda(\bar{y}) < k \end{cases}$

$E_{\theta_0}[\phi(\bar{y})] = P_{\theta_0}(\lambda(\bar{y}) > k) = \int_0^{r_1} f(y)\,dy + \int_{r_2}^{\infty} f(y)\,dy = \alpha$

$P(\text{type II error}) = P_{\theta_1}(\phi(\bar{y}) = 0) = P_{\theta_1}(\lambda(\bar{y}) < k) = P_{\theta_1}(r_1 < y < r_2)$

$= \int_{r_1}^{r_2} \frac{e^{-\sqrt{y}}}{2\sqrt{y}}\,dy = \int_{\sqrt{r_1}}^{\sqrt{r_2}} e^{-t}\,dt = -e^{-t}\Big|_{\sqrt{r_1}}^{\sqrt{r_2}} = \left(e^{-\sqrt{r_1}} - e^{-\sqrt{r_2}}\right)$

Q9 $X_i \sim N(0, \sigma^2)$

$f_{\theta_0}(\bar{x}) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(\frac{-\Sigma x_i^2}{2\sigma_0}\right)$

$\lambda(\bar{x}) = \frac{f_{\theta_1}}{f_{\theta_0}}(\bar{x}) = \frac{\sigma_0}{\sigma_1} \exp\left(\sum_i x_i^2 \left(\frac{1}{2\sigma_0} - \frac{1}{2\sigma_1}\right)\right)$; $\sigma_0 < \sigma_1 \Rightarrow \frac{1}{2\sigma_1} < \frac{1}{2\sigma_0} \Rightarrow \frac{1}{2\sigma_0} - \frac{1}{2\sigma_1} > 0$

$\therefore \lambda(\bar{x})$ inc. w/ $\sum_i x_i^2$

$\therefore \phi = \begin{cases} 1 ; & \lambda(\bar{x}) > c \\ 0 ; & \lambda(\bar{x}) < c \end{cases}$; $E_{\theta_0}[\phi(\bar{x})] = P_{\theta_0}(\phi(\bar{x}) = 1) = P_{\theta_0}(\lambda(\bar{x}) > c)$

$= P_{\theta_0}\left(\Sigma x_i^2 \ge \frac{2\sigma_0\sigma_1}{\sigma_1 - \sigma_0} \ln\left(\frac{\sigma_1}{\sigma_0} c\right)\right) = \alpha$

$\frac{2\sigma_0\sigma_1}{\sigma_1 - \sigma_0} \ln\left(\frac{\sigma_1}{\sigma_0} c\right) = \Sigma x_i^2$

Page 2

Lecture 1

Tuesday, July 18, 2017 8:36 AM

Data Representation

ID	Value
1	x_1
2	x_2
$\vdots$	$\vdots$
i	x_i
$\vdots$	$\vdots$
N	x_N

or

Value	frequency
y_1	f_1
y_2	f_2
$\vdots$	$\vdots$
y_m	f_m
y_M	f_M

$x_i \in \{y_1, y_2, \dots, y_M\} \; \forall i = 1, 2, \dots, N \quad (M < N)$

$$f_m = \sum_{i=1}^{N} 1_{\{x_i = y_m\}}$$

$$1_{\{x_i = y_m\}} = \begin{cases} 1 & \text{if } x_i = y_m \\ 0 & \text{otherwise} \end{cases}$$

Relative frequency $\lambda_m = \dfrac{f_m}{\sum_m f_m} = \dfrac{f_m}{N}$

Sample Mean $\bar{x} = \dfrac{\sum_{i=1}^{N} x_i}{N} = \dfrac{\sum_{j=1}^{M} y_m f_m}{N}$

(Assume w/o loss of generality, $x_1 \le x_2 \le \dots \le x_N$)

Median m

Value s.t. 50% of entries are $\le m$ and 50% are $\ge m$

if N is odd, $m = x_{\left(\frac{N+1}{2}\right)}$

if N is even, $m = \left(\dfrac{x_{N/2} + x_{\left(\frac{N}{2}+1\right)}}{2}\right)$ → Any # in $\left(x_{\frac{N}{2}}, x_{\left(\frac{N}{2}+1\right)}\right)$ satisfies the definition

↳ To define uniquely, we pick their mean

Percentile $p \in [0,1]$

p - percentile – x_p

$100p$ are $\le x_p$

& $100(1-p)$ are $\ge x_p$ → Median is $p = 0.5$

Quartile – $p = 0.25, 0.5, 0.75, 1$

Problem find x^* s.t., minimize –

1. $\sum |x_i - x|^2$
2. $\sum |x_i - x|$

1. $f(x) = \sum_i |x_i - x|^2 = \sum_i (x_i^2 - 2x_i x + x^2)$

$$= \sum_i x_i^2 - 2x \sum x_i + nx^2$$

$f'(x) = 0 - 2\left(\sum_i x_i\right) + 2nx$; $f' = 0 \Rightarrow x = \dfrac{\sum_i x_i}{n}$

$f''(x) = 2n > 0 \Rightarrow f'$ is minima

$\therefore \; x = \dfrac{\sum_i x_i}{n} = \bar{x}$ is the minima

2. $f(x) = \sum_i |x_i - x|$; Arrang x_i so that $x_i \le x_{i+1}$

Allowing f values from data set,

$$f(x_j) = \sum_i |x_i - x_j| = \sum_{i=1}^{j} (x_j - x_i) + \sum_{j+1}^{n} (x_i - x_j)$$

$$f(x_j) = j x_j - \sum_1^j x_i + \sum_{j+1}^{n} x_i + (j-n) x_j = (2j-n) x_j - 2\sum_1^j x_i + \sum_1^n x_i$$

Minima when $f(x_j) < f(x_{j+1})$ & $f(x_j) < f(x_{j-1})$

Lecture 2

Thursday, July 20, 2017 9:55 AM

Sample Variance

$$s^2 = \frac{1}{n-1}\sum_1^n (x_i - \bar{x})^2$$

$n-1$, not n
$\because$ $\exists$ unbiased estimates (to do later)

Data set 1

Value	freq
1	50
-1	50

$\bar{x}_1 = 0$

$s_1^2 = \frac{100}{99}$

Data set 2

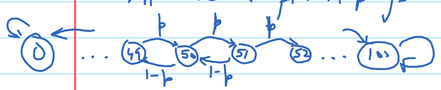
Value	freq
100	50
-100	50

$\bar{x}_2 = 0$

$s_2^2 = \frac{10^6}{99}$

$\frac{100^2 \times 100}{99}$

x_n = Wealth after n plays



P_w = probability that you eventually win all, starting from wealth w

$P_{100} = 1$, $P_0 = 0$

$P_w = p P_{w+1} + (1-p) P_{w-1}$

Chebyshev's Inequality

$S_k = \{ i : |x_i - \bar{x}| < ks \}$; $s = +\sqrt{s^2}$ (std. deviation)

Theorem: $\frac{|S_k|}{n} \geq 1 - \frac{1}{k^2}$

$\frac{|S_k^c|}{n} \leq \frac{1}{k^2}$

Proof $s^2 = \frac{1}{n-1}\sum_1^n (x_i - \bar{x})^2 = \frac{1}{n-1}\left(\sum_{i \in S_k}(x_i - \bar{x})^2 + \sum_{i \in S_k^c}(x_i - \bar{x})^2\right)$

$\geq \frac{1}{n-1}\left(\sum_{i \in S_k^c}(x_i - \bar{x})^2\right) \geq \frac{1}{n-1}\left(\sum_{i \in S_k^c} s^2 k^2\right) = \frac{1}{n-1}|S_k^c| s^2 k^2$

Lecture 4

Monday, July 24, 2017 11:39 AM

1. Chebyshev's Inequality

$S = +\sqrt{S^2}$ = standard deviation

$S_k = \{i : |x_i - \bar{x}| < ks\}$

$\boxed{\frac{|S_k|}{n} > 1 - \frac{1}{k^2}}$ or $\boxed{\frac{|S_k^c|}{n} \le \frac{1}{k^2}}$

2. One sided Chebyshev's Inequality

$\tilde{S}_k = \{i : x_i - \bar{x} \ge ks\}$; s = standard deviation

then $\frac{|\tilde{S}_k|}{n} \le \frac{1}{1+k^2}$

Proof: $\frac{\sum (x_i - \bar{x})^2}{n-1} = s^2$; $y_i = x_i - \bar{x}$

for any $b > 0$, $\sum_i (y_i + b)^2 = \sum_{i \in \tilde{S}_k} (y_i + b)^2 + \sum_{i \in \tilde{S}_k^c} (y_i + b)^2$

$\ge \sum_{i \in \tilde{S}_k} (ks + b)^2 = |\tilde{S}_k| (ks+b)^2$

Also, $\sum_i (y_i + b)^2 = \sum_i (y_i^2 + 2y_i b + b^2)$

$= \sum_i y_i^2 + 2b \sum_i y_i + nb^2 = (n-1)s^2 + 0 + nb^2$

$\Rightarrow (n-1)s^2 + nb^2 \ge |\tilde{S}_k| (ks+b)^2$

$\Rightarrow |\tilde{S}_k| \le \frac{(n-1)s^2 + nb^2}{(ks+b)^2} \Rightarrow \frac{|\tilde{S}_k|}{n} \le \frac{b^2 + \frac{n-1}{n}s^2}{(ks+b)^2} < \frac{b^2+s^2}{(ks+b)^2} = g(b)$

$2b = g'(ks+b)^2 + 2\cancel{(ks+b)} \times \frac{b^2+s^2}{(ks+b)^{\cancel{2}}}$

Put $g' = 0 \Rightarrow \cancel{2}b = \cancel{2}\frac{b^2+s^2}{ks+b} \Rightarrow \cancel{b^2} + bks = \cancel{b^2} + s^2$

$\Rightarrow b = s/k$

$\therefore \frac{|\tilde{S}_k|}{n} < \frac{s^2/k^2 + s^2}{s^2(k+\frac{1}{k})^2} = \frac{1+k^2}{k^2} \times \frac{1}{\frac{1}{k^2}(k^2+1)^2} = \frac{1}{1+k^2} \Rightarrow \frac{|\tilde{S}_k|}{n} = \frac{1}{1+k^2}$

7/25/2017 3:07 AM

Correlation:

Id.	Val1	Val2
1	x_1	y_1
2	x_2	y_2
⋮	⋮	⋮
n	x_n	y_n

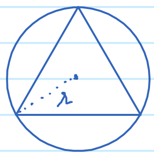
$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{(n-1) s_x s_y}$

Ex $p_k = P(\text{family has 'k' kids})$; $k = 1, 2, 3, \ldots, 10$

Probability: $P(\text{Event}) = \dfrac{\#\text{ of favourable outcomes}}{\text{Total }\#\text{ of outcomes}}$ } No metric of weighting More likely outcomes

$$P(\text{Event}) = \lim_{n \to \infty} \left(\frac{\#\text{ of time event happens in 'n' trials}}{n} \right)$$ } Doesn't converge

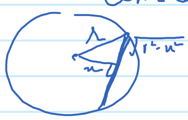
Ex



Circle of radius 'r'

$P(\text{randomly drawn chord has length} > \sqrt{3}\, r) = ?$

Correlate each pt. in circle to a unique chord using mid-point



$l = 2\sqrt{r^2 - x^2}$ $\quad l \geq \sqrt{3}\, r$

$\Leftrightarrow 4(r^2 - x^2) \geq 3r^2$

$\cancel{4} r^2 - 4x^2 \geq \cancel{3r^2}$

$r^2 \geq 4x^2$

$\Leftrightarrow \left(x \leq \frac{r}{2} \right)$

$\therefore P(\_\_) = \dfrac{\pi x^2}{\pi r^2} = \dfrac{\pi \frac{r^2}{4}}{\pi r^2} = \dfrac{1}{4}$

7/27/2017 10:18 AM

"The spirit of Probability"

Experiment Ω – set of all possible outcomes
– aka Sample Space

Event – set of all favourable outcomes
– Event $\subseteq \Omega$

'F' is a collection of events we are interested in

Lecture 7

Monday, July 31, 2017 11:43 AM

Probability

Ω is a collection of all outcomes

$A \subseteq \Omega$ is an event

'F' a collection of events

'F' satisfies
- (i) If $A \in F$ then $A^c \in F$
- (ii) If $A_1, A_2, \ldots \in F$ then $\bigcup_{i=1}^{\infty} A_i \in F$ – Countably infinite
- (iii) $\Omega \in F$

For F satisfying above three, we call F a σ-field

properties of F

- (i) $\phi \in F$
- (ii) $A_1, A_2, \ldots, A_n \in F$, then $\bigcup_{i=1}^{n} A_i \in F$
- (iii) If $A_1, A_2 \in F$, then $A_1 \cap A_2 \in F$
- (iv) (iii) is true for countably infinite sets
- (v) If $A, B \in F$, then $A-B$, $B-A$, $A \Delta B \in F$

The pair (Ω, F) is called a measurable space

Probability Measure (P): $P: F \to [0,1]$

- i) $P(\Omega) = 1$
- ii) $A_1, A_2, \ldots \in F$ are mutually disjoint, $P(\bigcup_i A_i) = \sum_i P(A_i)$
- iii) $P(A^c) = 1 - P(A)$
- iv) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Continuity from below of P

We say that $\lim_{h \uparrow 0} f(x+h) = y_0$

if $\forall$ sequences $\{h_n\}$ s.t. $h_n < 0\ \forall n$ and $\lim_{n\to\infty} h_n = 0$, $f(x+h_n) \to y_0$

Def. of continuity from below

For P, we need to
- i) Make a sequence of sets
- ii) Find limit of this sequence of sets

8/1/2017 8:39 AM

Assignment: Given a collection of events $A = \{A_1, A_2, \ldots, A_n\} \subseteq F$. Let A_σ be a σ-field that contains A. Then show that

Assignment: Given a collection of events $A = \{A_1, A_2, \dots, A_n\} \subseteq F$
Let A_r be a σ-field that contains A. Then show that
$\exists$ a unique smallest σ-field containing A, $\sigma(A)$

- Continuity from below
sequence A_n's; $A_n \subseteq A_{n+1}$, $A_\infty = \bigcup_{i=1}^{\infty} A_i$

$$P\left(\lim_{N\uparrow\infty} \bigcup_{i=1}^{N} A_i\right) = \lim_{n\uparrow\infty} P(A_n)$$

def. sequence B, $B_1 = A_1$
$B_n = A_n - A_{n-1}$

Claim: $B_i \cap B_j = \phi \ \forall i \neq j$
w/o loss of generality, $i < j$

$$B_i \cap B_j = (A_i - A_{i-1}) \cap (A_j - A_{j-1})$$

$(A_i - A_{i-1}) \subseteq A_{j-1}$; $A_{j-1} \cap (A_j - A_{j-1}) = \phi$
$\Rightarrow$ claim proven

$\bigcup_{i=1}^{n} B_i = A_n$; $\Rightarrow P(A_n) = \sum_{i=1}^{n} P(B_i) \quad \forall n$

$$\lim_{n\uparrow\infty} P(A_n) = \lim_{n\uparrow\infty} \sum_{1}^{n} P(B_i) = \sum_{1}^{\infty} P(B_i) = P\left(\bigcup_i B_i\right)$$

$$\Rightarrow \lim_{n\uparrow\infty} P(A_n) = P\left(\lim_{n\to\infty} \bigcup_{i=1}^{n} B_i\right) = P\left(\lim_{n\uparrow\infty} A_n\right) = P\left(\lim_{n\uparrow\infty} \bigcup_{i=1}^{n} A_i\right)$$

Continuity from above: $A_n \supseteq A_{n+1}$; $A_\infty = \bigcap_{n=1}^{\infty} A_n$

$$\lim_{n\uparrow\infty} P(A_n) = P\left(\lim_{n\uparrow\infty} A_n\right)$$

Ex $\Omega = \{H, T\}$
$F = \{\phi, \{H\}, \{T\}, \Omega\}$

$P(\{H\}) = p$; $P(\{T\}) = 1-p$

Probability Measure Def: let $D \in F$ s.t. $P(D) > 0$
$P_{|D} : F \to [0,1]$
s.t. $P_{|D}(A) = \dfrac{P(A \cap D)}{P(D)} \ \forall A \in F$

$\rightarrow$ Show that $P_{|D}$ is a valid probability measure
i) $P_{|D}(\Omega) = 1$ and $P_{|D}(\phi) = 0$

Lecture 9

Thursday, August 3, 2017 9:42 AM

Baye's Theorem If $A, D \in F$ s.t. $P(A), P(D) > 0$ then

$$P(A|D) = \frac{P(D|A)\cdot P(A)}{P(D)}$$

Independence

(of 2 events) Events A_1 & A_2 are independent if $P(A_1 \cap A_2) = P(A_1)\cdot P(A_2)$

notation: $A_1 \perp\!\!\!\perp A_2$

(of a collection of events) The collection A on 'n' events is independent if $\forall S \subseteq \{1,2,3,\dots,n\}$

$$P\left(\bigcap_{i\in S} A_i\right) = \prod_{i\in S} P(A_i)$$

Prove $P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$ $\not\Leftrightarrow$ $P(A_i \cap A_j) = P(A_i)\cdot P(A_j)\ \forall i \neq j$ $\longrightarrow$ both ways not true

(for independence of collections) $A = \{A_1,\dots,A_n\} \subseteq F$, $B = \{B_1,\dots,B_m\} \subseteq F$

Collection $A \perp\!\!\!\perp B$ if $A_i \perp\!\!\!\perp B_j\ \forall i \in \{1,\dots,n\}$ & $j \in \{1,\dots,m\}$

2 dice thro

$\frac{1}{3}$ (1x 3x) (1x 2x) $\frac{1}{3}$

(2x 3x) $\frac{1}{3}$

Random Variable (Ω, F, P) $X: \Omega \to \mathbb{R}$

$X(\omega) \in \mathbb{R}\ \forall \omega \in \Omega$

8/7/2017 11:39 AM

We take $\Omega = \mathbb{R}$ (generally)

We typically use Borel σ-algebra $B(\mathbb{R})$

$\downarrow$ Smallest σ-field on $\mathbb{R}$ containing $\{(-\infty, x] : x \in \mathbb{R}\}$

$\therefore X : (\Omega, F, P) \to (\mathbb{R}, B(\mathbb{R}), P_x)$

$P_x : B(\mathbb{R}) \to [0,1]$; $P_x(B) = P(\{\omega : X(\omega) \in B(\mathbb{R})\})\ \forall B \in B(\mathbb{R})$

Ex Coin toss (Ω, F, P)

$\Omega = \{H, T\}$

F = Power set of Ω

$P(\{H\}) = p$, $P(\{T\}) = 1-p$

Define a random variable X

$X(H) = -1$; $X(T) = +1$

$\{\omega \cdot X(\omega) \in B(\mathbb{R})\} = X^{-1}(B)$

$$P_x(B) = P(X^{-1}(B))$$

$$P_x((-\infty, x]) = \text{Prob}\{X \in (-\infty, x]\} = \text{Prob}\{X \le x\}$$

$$= F_x(x) \to \text{distribution fn.}$$

Imp Theory

(Back to ex)

$$\text{Prob}\{X \le x\} = \begin{cases} 0 ; & x < -1 \\ p ; & -1 \le x < 1 \end{cases}$$

$\left. 1 \;;\; 1 \le x \right.$

HW $F_X(x)$ is cont. from right w/ left hand limit (cadlag)

Lecture 11

Tuesday, August 8, 2017 8:37 AM

{ **Assignment**
— Show that P_X is a valid prob. measure }

Recap (of this)
$X \sim F_X(\,)$
1. $F_X(-\infty)=0$, $F_X(\infty)=1$
2. F_X is monotone, non-decreasing
3. F_X is cadlag

Types of random variables
Continous — F_X is cont. fn.
Discrete — F_X is a staircase fn.

Density & prob. mass fn.
Let $X \sim F_X$ is continous
$$\int_{-\infty}^{x} F_X(u)\,du = f_X(x)$$

Properties of F_X
(i) $F_X(-\infty)=0$ & $F_X(+\infty)=1$
(ii) F_X is a monotone non-decreasing fn.
(iii) F_X is right continous & left hand limit exists

proofs
(i) $F_X(-\infty) = P(x \in (-\infty,-\infty))$; $F_X(\infty) = P(x \in (-\infty,\infty))$
$= P(\phi) = 0$; $= P(\Omega) = 1$

(ii) $X^{-1}((-\infty, x_1]) \subseteq X^{-1}((-\infty, x_2])$ when $x_1 \leq x_2$

(iii) T.P. $\lim_{n \uparrow \infty} F_X(x_n) = F_X(\lim_{n\uparrow\infty} x_n)$

x_n is a monotone dec. sq. approaching x i.e. $\lim x_n = x$

$F_X(x_n) = P(X^{-1}((-\infty, x_n]))$
$X^{-1}((-\infty, x_n]) = A_n$
$A_{n+1} \subseteq A_n \ \forall n$

$$\begin{aligned}\lim_{n\uparrow\infty} P(A_n) &= P(\lim_{n\uparrow\infty} A_n) \\ &= P(\bigcap_1^{\infty} A_n) = P_X(\{\lambda : X(\omega)=\lambda, \omega \in \bigcap_1^{\infty} A_n\}) \\ &= P_X(\bigcap_{n=1}^{\infty}\{\lambda : X(\omega)=\lambda, \omega \in A_n\}) \\ &= P_X(\bigcap_{n=1}^{\infty} (-\infty, x_n]) \\ &= P_X((-\infty, x]) \\ &= F_X(x) = F_X(\lim_{n\uparrow\infty} x)\end{aligned}$$

Types of Random Variables
1. Continous — if $F_X(x)$ is continous
2. Discrete — if $F_X(x)$ is a staircase fn.

Any random variable can be expresses as a linear combination of continous & discrete random variable

8/10/2017 9:39 AM

For $X \sim F_X$ being discrete,

Let $\varepsilon \in (0,1)$
At most n steps of size ε or more
$\therefore \varepsilon_n = \frac{1}{n}$

$D_n = \{r : \text{jump size} \geq \frac{1}{n} \text{ at } \lambda\}$

$|D_n| \leq n$; Each D_n is countable

Let D be the set of all jump epochs of F_X — points of jumps

$D = \bigcup_{n=1}^{\infty} D_n$ — proof
[Suppose false, $\Rightarrow \exists\, r \in D$ & $r \notin \bigcup_i D_i$
Assume jump at λ is $j > 0$
Let $\frac{1}{\alpha} = \lfloor j \rfloor + 2$. So $\lambda \in D_\alpha \rightarrow$ Contradiction]

Prove: Countable union of countable sets is countable

So D is countable
$\therefore$ # of jump epochs is atmost countably many

$$p_k = F_X(x_k) - F_X(x_k^-)$$

$p_k = P\{x = x_k\}$ prob. mass fn.

$$f_X(x) = \sum_{k=1}^{\infty} p_k \underbrace{\delta(x - x_k)}_{\text{Dirac Delta}}$$

Show that $\int_{-\infty}^{x} f_X(u)\,du = F_X(x)$

Moments of Random Variables

'k'th moment of x

$E[x^k] = \int_{-\infty}^{\infty} x^k f_X(x)\,dx$; whenever the integral is well defined

→ first moment is called expectation

→ Absolute moments

$$E[|x|^k] = \int_{-\infty}^{\infty} |x|^k f_X(x)\,dx$$

→ Central moments

$$E[(x - E_X)^k] = \int_{-\infty}^{\infty} (x - E_X)^k f_X(x)\,dx$$

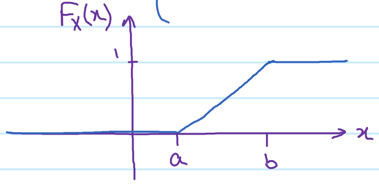
Lecture 14

Thursday, August 17, 2017 9:37 AM

Common continous Random Variable (r.v.)

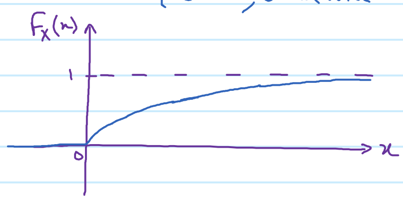
1. Uniform r.v. (a,b) $b>a$

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



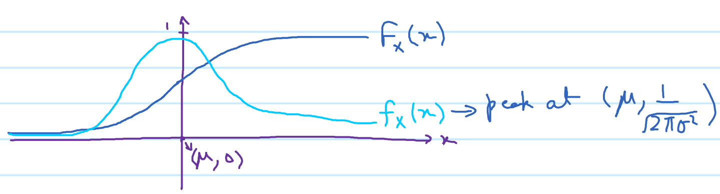
2. Exponential r.v. $(\lambda \in \mathbb{R}^+)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & ; \text{for } x \geq 0 \\ 0 & ; \text{otherwise} \end{cases}$$



3. Gaussian r.v. (aka. Normal r.v.)
   (μ, σ^2 are parameters)

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-(x-\mu)^2/2\sigma^2} \quad \forall\, x \in \mathbb{R}$$



ERFC fn. is Φ

$$\Phi(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}}\, e^{-u^2/2}\, du$$

$\downarrow$

[Prob. that a Gaussian r.v. w/ $\mu=0, \sigma^2=1$ takes values $\geq x$]

$\to$ Reading Assgn. $\to$ CH5 of Rohatgi & Saleh (or Sheldon Ross?)

Functions of R.V.

$X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$

$g: \mathbb{R} \to \mathbb{R}$

$g(X)$ as a random variable on (Ω, F)

$g(X): \Omega \to \mathbb{R}$

$g(\underbrace{X(\omega)}_{\text{read as a \#, say } x}) = g(x)$

Ex $Y = g(X)$
$X \sim F_X$ assumed continous

1. Find F_Y

$$F_Y(y) = P(Y \le y)$$
$$= P(g(X) \le y)$$

- for invertible $g(X)$, let $g(x_1) = y$

$$g(X) \le y \Leftrightarrow x \le x_1 \ (\text{or} \ge x_1)$$

- for non invertible g

Lecture 16

Tuesday, August 22, 2017 8:40 AM

Recap

- Fn. of random variable
  $g : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$
  Borel measurable
  $Y = g(X)$
- $f_Y(y) = \sum_{k=1}^{\infty} \frac{f_X(x_k)}{|g'(x_k)|}$; x_k(s) are roots of $y = g(x)$

- Expectation
  $E[x] = \int_{-\infty}^{\infty} x f_X(x)\,dx$ - continous
  $= \sum_{k=1}^{\infty} a_k p_k = \sum_{k=1}^{\infty} a_k P(X = a_k)$

Properties: (Assignment)

1. If f_X is symmetric ($f_X(-x) = f_X(x)$)
   then $E[x] = 0$

2. If f_X is symmetric about 'u'
   i.e. $f(u + x) = f(u - x)$
   then $E[x] = u$

3. For bounded r.v., i.e.
   $|x| \le b \Rightarrow |X(\omega)| \le b \ \forall \omega$
   $f_X(x) = 0 \ \forall |x| > b$
   $F_X(x) = \begin{cases} 0 & \forall x < -b \\ 1 & \forall x > b \end{cases}$
   $E[x] = b - \int_{-b}^{b} F_X(x)\,dx$

4. Expectation of a fn. of r.v.
   $Y = g(X)$; for $E[Y]$
   (i) $E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\,dy$ (or)
   (ii) $E[Y] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$
   (iii) $E[x^k] = \int_{-\infty}^{\infty} x^k f_X(x)\,dx$

TP: $\sum_{j=1}^{\infty} b_j P(Y = b_j) = \sum_{k=1}^{\infty} g(a_k) P(X = a_k)$ (discrete)

$I_j = \{k : g(a_k) = b_j\}$
$P(Y = b_j) = \sum_{k \in I_j} P(X = a_k)$

$\Rightarrow \sum_j b_j P(Y = b_j) = E(Y) = \sum_j b_j \sum_{k \in I_j} P(X = a_k) = \sum_j \sum_{k \in I_j} b_j P(X = a_k)$

$\Rightarrow E(Y) = \sum_j \sum_{k \in I_j} g(a_k) P(X = a_k)$; I_j is a partition of Domain, so

$\Rightarrow E(Y) = \sum_k g(a_k) P(X = a_k)$

(iv) Moment Generating fn.
   for r.v. X, mgf is def. as
   $E[e^{sX}] = \int_{-\infty}^{\infty} e^{sx} f_X(x)\,dx$
   $(s > 0)$
   $\psi(s) = E[e^{sX}]$
   $\psi(s)$ is well-defined for $s \in (0, \delta)$
   $\psi'(s) = \frac{d}{ds} \int_{-\infty}^{\infty} (\,) = \int_{-\infty}^{\infty} \frac{d}{ds} (\,) = \int_{-\infty}^{\infty} x e^{sx} f_X(x)\,dx$

   $\psi'(s)\big|_{s=0} = \int_{-\infty}^{\infty} x f_X(x)\,dx$

   $\psi''(s)\big|_{s=0} = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx$

Ex $f_X(x) = \begin{cases} \lambda e^{-\lambda x}; & x \geq 0 \\ 0 & ; x < 0 \end{cases} \quad (\lambda > 0)$

$$E[e^{sX}] = \int_{-\infty}^{\infty} e^{sx} f_X(x)\, dx = \int_0^{\infty} e^{sx} \lambda e^{-\lambda x} dx = \frac{\lambda}{s-\lambda}\int_0^{\infty} e^{(s-\lambda)x} d(s-\lambda)x = \frac{\lambda}{s-\lambda} e^{(s-\lambda)x}\Big|_0^{\infty}$$

$$\underset{s\in(0,\lambda)}{E[e^{sX}]} = \frac{\lambda}{s-\lambda}[0-1] = \frac{\lambda}{\lambda - s} = \frac{1}{1-\left(\frac{s}{\lambda}\right)}$$

Ex $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$$E[e^{sX}] = \int_{-\infty}^{\infty} e^{sx} f_X(x)\, dx = \int_{-\infty}^{\infty} e^{sx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2+2sx-s^2}{2}\right) e^{s^2/2} dx$$

$$= \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}} e^{s^2/2}\right) \exp\left(\frac{-1}{2}(x-s)^2\right) dx = e^{s^2/2} \times \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-s)^2}{2}} dx\right]$$

Gaussian density w/ mean $=1$
so $\int_{-\infty}^{\infty}(\ ) = 1$

$\Rightarrow E[e^{sX}] = e^{s^2/2}$

Lecture 17

Thursday, August 24, 2017 9:38 AM

Characteristic fn.

$\phi_X(\omega) = E[e^{j\omega X}] = \int_{-\infty}^{\infty} e^{j\omega x} f_X(x)\, dx$ } Fourier Transform

$-\psi_X(t)$ i.e. m.g.f. doesn't always exist. $\phi_X(\omega)$ always exists.

$-\phi_x(\omega)$ is an equivalent representation of $f_X(x)$

2 Random Variables

(Ω, F, P) ; X & Y are 2 r.v. def. on (Ω, F)

w/ 2 r.v., $\{X \le x \ \& \ Y \le y\}$ is a possible event

↓ written as

$\{\omega : X(\omega) \le x\} \cap \{\omega : Y(\omega) \le y\}$

$F_{XY}(x,y) = P(X \le x; Y \le y)$

Joint distribution fn.

$F_{XY}(-\infty, -\infty) = 0$

$F_{XY}(-\infty, y) = 0$

$F_{XY}(+\infty, +\infty) = 1$

$F_{XY}(+\infty, y) = F_Y(y) \rightarrow$ Marginal distribution fn.

8/29/2017 8:43 AM

Recap:

Marginals $F_X(x) = F_X(x, +\inf.) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{XY}(u,v)\, dv\, du$

Marginal Density $f_X(x) = \frac{dF_X(x)}{dx} = \int F_{XY}$

Conditional prob. Let A be an event w/ $P(A) > 0$

$$F_Y(y|A) = P(Y \le y | A)$$
$$= \frac{P(Y \le y, A)}{P(A)}$$

Event $A = \{X \le x\}$

$$F_Y(y | X \le x) = \frac{P(Y \le y, X \le x)}{P(X \le x)}$$
$$= \frac{F_{XY}(x,y)}{F_X(x)}$$

$A = \{X > x\}$

$F_Y(y | X > x) = \frac{P(Y \le y, X > x)}{}$

$$P(X>x)$$

$$= \frac{F_Y(y) - F_{XY}(x,y)}{1-F_X(x)}$$

$$\underbrace{f_Y}_{\text{conditional density}}(y|A) = \frac{\partial}{\partial y}(F_Y(y|A))$$

for $A = \{X \le x\}$

$$f_Y(y|A) = \frac{1}{F_X(x)} \frac{\partial}{\partial y} F_{XY}(x,y)$$

$$= \frac{1}{F_X(x)} \frac{\partial}{\partial y}\left[\int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(u,v)\,dv\,du\right]$$

$$= \frac{1}{F_X(x)}\left[\int_{-\infty}^{x} \underbrace{\frac{\partial}{\partial y}\int_{-\infty}^{y} f_{XY}(u,v)\,dv}_{f_{XY}(u,y)}\,du\right]$$

$$= \frac{1}{F_X(x)}\left[\int_{-\infty}^{x} f_{XY}(u,y)\,du\right]$$

for $A = \{X = x\}$

$$F_Y(y|X=x) = \lim_{\Delta x \downarrow 0} F_Y(y \mid x < X \le x+\Delta x)$$

$$= \lim_{\Delta x \downarrow 0} \frac{P(Y \le y,\ x < X \le x+\Delta x)/\Delta x}{\left(P(x < X \le x+\Delta x)/\Delta x\right) \to f_X(x)}$$

$$= \frac{\frac{\partial}{\partial x} F_{XY}(x,y)}{f_X(x)} = \frac{\int_{-\infty}^{y} f_X(x,v)\,dv}{f_X(x)}$$

Law of total Probability

$$f_Y(y) = \int_{-\infty}^{\infty} f_Y(y|X=x)\, f_X(x)\,dx$$

$\exists\ \Omega$

Let $A_1, A_2, A_3, \ldots, A_n$ be a partition of Ω

Then $P(B) = \sum_{k=1}^{n} P(B|A_k)P(A_k)$

Baye's Rule $f_X(x|Y=y) = \dfrac{f_Y(y|X=x)\, f_X(x)}{f_Y(y)}$

Lecture 19

Monday, September 4, 2017 11:42 AM

- Markov Inequality

$$P(|X| > \varepsilon) \le \frac{E|x|^k}{\varepsilon^k} \quad ; \ \varepsilon > 0, k \ge 1$$

- Chebyshev's Inequality

$$P(|X - E_X| > k_J) \le \frac{1}{k^2}$$

- Jensen's Inequality

$g : \mathbb{R} \to \mathbb{R}$ is a concave fn.

$$E[g(x)] \le g(E_X)$$

Independence of R.V.

R.V. X & Y are indep. if

$$F_{XY}(x,y) = F_X(x) \cdot F_Y(y) \ \forall \ x, y \in \mathbb{R}$$

iid → indep. & identically distributed

X & Y are iid if $X \perp\!\!\!\perp Y$ & $F_X(u) = F_Y(u) \ \forall \ u \in \mathbb{R}$

For A & B indep.

$$P(A|B) = P(A)$$

$$\left\{ = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\cancel{P(B)}}{\cancel{P(B)}} \right\}$$

- Show that

$F_Y(y | X \le x) = F_Y(y)$

$F_Y(y | X \in B) = F_Y(y)$; B is a Borel set

(when $X \perp\!\!\!\perp Y$)

$E[XY] = E[X]\,E[Y]$ (*does not imply indep.)

Q/ If $E[X^r Y^s] = E[X^r]E[Y^s] \ \forall \ r, s$, does this imply indep.?

- X & Y are uncorrelated if $E[XY] = E[X]E[Y]$
- X & Y are orthogonal if $E[(X - E_X)(Y - E_Y)] = 0$

Define: Correlation coefficient

$$\rho_{XY} = \frac{E[(X - E_X)(Y - E_Y)]}{\sigma_X \sigma_Y}$$

2 fn. of 2 R.V.

$$Z = g(X,Y), \quad W = h(X,Y)$$

- find joint density / distribution of Z & W

$$f_{ZW}(z, w) = \sum_{k=1}^{\infty} \frac{f_{XY}(x_k, y_k)}{|J(x_k, y_k)|}$$

$$|J(x_k, y_k)|$$

where $J(x_k, y_k) = \begin{vmatrix} \frac{\partial}{\partial x} g(x,y) & \frac{\partial}{\partial y} g(x,y) \\ \frac{\partial}{\partial x} h(x,y) & \frac{\partial}{\partial y} h(x,y) \end{vmatrix}$

where (x_k, y_k) are roots of
$g(x,y) = z$ & $h(x,y) = w$

ex $Z = aX + bY$ & $W = cX + dY$

$$J(x,y) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = (ad - bc)$$

$z = ax + by$ & $w = cx + dy$

$$x = \frac{dz - bw}{ad - bc} \; ; \; y = \frac{aw - cz}{ad - bc}$$

$$f_{ZW}(z,w) = \frac{f_{XY}\left(\frac{dz-bw}{ad-bc}, \frac{aw-cz}{ad-bc}\right)}{|ad-bc|}$$

Ex $Z = +\sqrt{X^2+Y^2}$ $\quad W = X/Y$

$$\frac{\partial g}{\partial x} = \frac{x}{\sqrt{x^2+y^2}} \qquad \frac{\partial g}{\partial y} = \frac{y}{\sqrt{x^2+y^2}}$$

$$\frac{\partial h}{\partial x} = \frac{1}{y} \qquad \frac{\partial h}{\partial y} = \frac{-x}{y^2}$$

$$J(x,y) = -\frac{x^2}{y^2 g} - \frac{1}{g} = -1\frac{\overset{g}{\cancel{(x^2+y^2)}}}{y^2 \cancel{g}} = -\frac{\sqrt{x^2+y^2}}{y^2}$$

$$x = wy \; ; \; z = \pm\sqrt{w^2+1}\, y$$

$$y_1 = \frac{z}{\sqrt{w^2+1}} \qquad x_1 = \frac{wz}{\sqrt{w^2+1}}$$

$$y_2 = -\frac{z}{\sqrt{w^2+1}} \qquad x_2 = -\frac{wz}{\sqrt{w^2+1}}$$

For n-Random Variables

$X_1, X_2, \ldots, X_n$

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = P(X_1 \le x_1, \ldots, X_n \le x_n)$$

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \frac{\partial^n}{\partial x_1 \partial x_2 \ldots \partial x_n} F_{X_1,\ldots X_n}(x_1,\ldots,x_n)$$

ex $f_{X_1,X_2,X_3}(x_1,x_2,x_3)$ given, find $f_{X_1,X_2}(x_1,x_2)$

$$f_{X_1,X_2}(x_1,x_2) = \int_{-\infty}^{\infty} f_{X_1,X_2,X_3}(x_1,x_2,x_3)\, dx_3$$

Lecture 20

Tuesday, September 5, 2017 8:39 AM

Fundamental Theorems of Probability Theory

1. Weak law of large numbers

Let $X_1, X_2, \ldots$ be i.i.d random variables w/ mean $E[X]$ then

$$P\left(\left|\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) - E[X]\right| > \varepsilon\right) \to 0 \text{ as } n \to \infty \quad \forall \varepsilon > 0$$

Let $Y_n = \frac{1}{n}\sum_{k=1}^{n}(X_k - E_X)$

$$E[Y_n] = \frac{1}{n}\sum_{k=1}^{n} \underbrace{E(X_k - E_X)}_{=0\ \forall k}$$

$$= 0 \ \forall n$$

$$\text{var}[Y_n] = E[(Y_n - E_{Y_n})^2] = E[Y_n^2]$$

$$= E\left[\left(\frac{1}{n}\sum_{k=1}^{n}\underbrace{(X_k - E_X)}_{Z_k}\right)^2\right]$$

$$= E\left[\frac{1}{n^2}\left(\sum_{k=1}^{n} Z_k\right)^2\right] = \frac{1}{n^2}E\left[\left(\sum_{k=1}^{n} Z_k\right)^2\right]$$

$$= \frac{1}{n^2}E\left[\sum_{k=1}^{n} Z_k^2 + \sum_{i=1}^{n}\sum_{j\neq i} Z_i Z_j\right] = \frac{1}{n^2}\left(\sum_{k=1}^{n} E[Z_k^2] + \underbrace{\sum_{i\neq j} E[Z_i Z_j]}_{=0}\right)$$

$$= \frac{1}{n^2}\sum_{k=1}^{n} E((X_k - E_X)^2) = \frac{1}{n^2}\sum_{k=1}^{n}\text{var}(X) = \frac{1}{n}\text{var}(X)$$

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - E_X\right| > \varepsilon\right) \leq \frac{\text{var}\left(\frac{1}{n}\sum_{k=1}^{n} X_k - E_X\right)}{\varepsilon^2} = \frac{\text{var}(X)}{n\varepsilon^2} \to 0 \text{ for } n \to \infty$$

2. Central Limit Theorem

$X_1, X_2, \ldots$ i.i.d s w/ mean μ & var σ^2

$E[X_k] = \mu \ \forall k$

$\text{var}[X_k] = \sigma^2 \ \forall k$

The distribution of $\dfrac{\sum_{k=1}^{n}(X_k - \mu)}{\sqrt{n}\,\sigma}$ approaches unit normal distribution as $n \to \infty$

unit normal dist. $N(0,1)$ or $G(0,1)$ is:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

9/7/2017 9:36 AM

Recap

- Convergence in Prob.

  A seq. $Y_1, Y_2, \ldots$ of r.v. converges in prob. to a r.v. Y if $P(|Y_n - Y| > \varepsilon) \to 0$ as $n \uparrow \infty$

- Weak law of large nos

  Let $Y_n = \frac{1}{n}\sum_k X_k$ and $Y = E_X$. Then, $Y_n \to Y$ in prob.

- Convergence in dist.

  A seq. $Y_1, Y_2, \ldots$ of r.v. converges in dist. to Y

if $\forall y \in \mathbb{R}$

$$F_{Y_n}(y) \to F_Y(y) \text{ as } n \uparrow \infty$$

- Central Limit Theorem

Let $Y_n = \dfrac{\sum_k (x_k - \mu_x)}{\sqrt{n}\,\sigma}$ and $f_Y = G(0,1)$

Then, $Y_n \to Y$ in dist. as $n \uparrow \infty$

Lecture 22

Tuesday, September 19, 2017 8:37 AM

- iid. samples $X_1, X_2, \ldots, X_n$

  $X \sim F_{\bar{\theta}}$, $\bar{\theta}$ are the parameters

Data analysis is finding that F & $\bar{\theta}$

- Given: $X_1, X_2, \ldots, X_n$

  find F

  Arrange in ascending order

  $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ s.t. $X_k \leq X_{k+1}$

$$F_n^*(x) = \begin{cases} 0 & \text{if } x < X_{(1)} \\ \frac{k}{n} & \text{if } X_{(k)} \leq x < X_{(k+1)} \text{ for } k=1,2,\ldots,n-1 \\ 1 & \text{if } X_{(n)} \leq x \end{cases}$$

Claim: $F_n^*(x)$ is a r.v. $\forall x \in \mathbb{R}$

$$P(F_n^*(x) = k/n) = {}^nC_k (F(x))^k (1-F(x))^{n-k}$$

$$E[F_n^*(x)] = \sum_{j=0}^{n} \frac{j}{n} {}^nC_k (F(x))^k (1-F(x))^{n-k}$$

$$= \frac{1}{n} E[B(n, F(x))] = \frac{1}{\cancel{n}} \times \cancel{n} \times F(x) = F(x)$$

$\Rightarrow E[F_n^*(x)] = F(x)$

$$\text{var}[F_n^*(x)] = \frac{1}{n} F(x)(1-F(x)) \leq \frac{1}{4n}$$

As $n \to \infty$, $\text{var}[F_n^*(x)] \to 0$

$\to P(|F_n^*(x) - F(x)| > \varepsilon) \to 0$ as $n \to \infty$ $\forall x, \forall \varepsilon > 0$

$\to$ Glivenko-Cantelli Lemma

Given. $X_1, X_2, \ldots, X_n$ & functional form of their dist.,

Find parameters of the dist.

Sample space $\to \mathbb{R}^n$

parameter space $\to \Theta$

$\psi : \Theta \to \mathbb{R}$

9/21/2017 9:46 AM

$F_{\bar{\theta}}$

- Assume functional form of $F_{\bar{\theta}}$ is known, but $\bar{\theta}$ isn't known

  $\mathcal{X}$-sample space, i.e. $\{x_1, x_2, x_3, \ldots\} \in \mathcal{X}$

- Θ parameter space, i.e. $\bar{\theta} \in \Theta$

Let $\psi : \Theta \to \mathbb{R}$

Aim: Estimate ψ based on observations $(X_1, X_2, \ldots, X_n)$

Estimate $\delta : \mathcal{X} \to \mathbb{R}$

Estimator

δ should not use $\bar{\theta}$ or any unknown parameter

δ is called a statistic function

Ex $X_1, X_2, \ldots, X_n$ outcomes of n coin tosses

$\mathcal{X} = \{0,1\}^n$; $\theta = p = P(X_k = 1)$

$\Theta = [0,1]$ Estimator δ

$$\delta(x_1, x_2, \ldots, x_n) = \frac{1}{n}\sum_{k=1}^{n} x_k$$

$\delta(X_1, \ldots, X_n) = X_1$

$$\delta(x_1, \ldots, x_n) = pX$$

Ex $X_1, X_2, \ldots, X_n$ be i.i.d. Gaussian w/ $X_k \sim \mathcal{G}(\mu, \sigma^2)$

$\mathcal{X} = \mathbb{R}^n$, $\Theta = \mathbb{R} \times \mathbb{R}^+$

$\psi(\mu, \sigma^2) = \mu + \sigma^2$

$$\delta(X_1, \ldots, X_n) = \frac{1}{n}\sum_k x_k + \frac{1}{n-1}\sum_k (x_k - \bar{x})^2$$

<u>Error fn.</u>

Mean sq error $E_\delta\left[\left(\psi(\bar{\theta}) - \delta(x_1, \ldots, x_n)\right)^2\right]$

Estimator δ_1 is better that δ_2 if

$$\max_{\bar{\theta} \in \Theta}\left(E_{\delta_1}(\bar{\theta})\right) \le \max_{\bar{\theta} \in \Theta}\left(E_{\delta_2}(\bar{\theta})\right)$$

Lecture 25

Monday, September 25, 2017 10:39 AM

- Statistic $T: \mathbb{R}^n \to \mathbb{R}^n$, fn. of known things

- Sufficient statistic
  $F_\theta(x_1,\dots,x_n \mid T=t)$ is not a fn. of θ, $\forall \theta \in \Theta$

- Factorization: T is sufficient iff
  $P_\theta(x_1,\dots,x_n) = h(x_1,\dots,x_n)\, g_\theta(T)$
  $f_\theta(x_1,\dots,x_n) = h(x_1,\dots,x_n)\, g_\theta(T)$

→ Read from Rohtagi

Examples to use factorization criteria

Ex1

$X_1, X_2, \dots, X_n \sim \exp(\lambda)$; $\lambda > 0$, iids

$$f_\lambda(x_1,\dots,x_n) = \prod_k \lambda e^{-\lambda x_k} = \lambda^n \exp(-\lambda \sum_k x_k) \quad \text{if } \min\{x_1,\dots,x_n\} \geq 0$$

$$= \lambda^n \exp(-\lambda \sum_k x_k)\, \mathbb{1}_{\{\min x_k \geq 0\}}$$

Set $h(x_1,\dots,x_n) = \mathbb{1}_{\{\min x_k \geq 0\}}$

set $g_\lambda(T(x_1,\dots,x_n)) = \lambda^n \exp(-\lambda \sum_k x_k)$

$\therefore T(x_1,\dots,x_n) = \sum_k x_k$ is sufficient

Ex2

$X_1,\dots,X_n \sim \mathcal{G}(\mu,\sigma^2)$, $f_{X_1}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \forall x \in \mathbb{R}$

$$f_{X_1,\dots,X_n}(x_1,\dots,x_n) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left(\frac{-1}{2\sigma^2} \sum_k (x_k-\mu)^2\right)$$

$$= \underbrace{\frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left(\frac{-\sum x_k^2}{2\sigma^2} + \frac{\mu \sum x_k}{\sigma^2} - \frac{n\mu^2}{2\sigma^2}\right)}$$

Set $h(x_1,\dots,x_n) = 1$, set $g(T(x_1,\dots,x_n)) =$ ____

$\therefore T(x_1,\dots,x_n) = \left(\sum_k x_k^2, \sum_k x_k\right)$ } Sufficient

Ex3

$X_1,\dots,X_n \sim \text{Uniform}\{1,2,\dots,N\}$

$P(X_k = \mu) = \frac{1}{N} \;\forall \mu \in \{1,2,\dots,N\}$; 0 otherwise

$$p_N(x_1,\dots,x_n) = \left(\frac{1}{N}\right)^n ; \quad \min x_k \geq 1 \;\&\; \max x_k \leq N \quad (x_k \in \mathbb{Z})$$

$$= \left(\frac{1}{N}\right)^n \mathbb{1}_{\min x_k \geq 1} \mathbb{1}_{\max x_k \leq N}$$

$h(x_1,\dots,x_n) = \mathbb{1}_{\min x_k \geq 1}$

$g_N(T) = \left(\frac{1}{N}\right)^n 1_{\max x_k \leq N}$

$T(x_1, \ldots, x_n) = \max\{x_k\}$] Sufficient

9/26/2017 8:34 AM

Minimal Sufficient Statistic T is said to be so if

1) T is sufficient
2) for any other statistic S, $\exists g$ s.t. $g(S) = T$

Minimality Tests

T is minimal if given any two data sets $\bar{x}$ & $\bar{y}$ such that $\left(T(\bar{x}) = T(\bar{y})\right)$ iff $\left(\frac{f_\theta(\bar{x})}{f_\theta(\bar{y})} = k(\bar{x}, \bar{y})\right)$

Ex $X_1, \ldots, X_n \sim$ Bernoulli(p)

$$P_p(x_1, \ldots, x_n) = p^{\Sigma x_k}(1-p)^{n-\Sigma x_k}$$

$$\frac{P_p(\bar{x})}{P_p(\bar{y})} = p^{\Sigma x_k - \Sigma y_k}(1-p)^{\Sigma y_k - \Sigma x_k} = 1 \left(\text{if } \Sigma x_k = \Sigma y_k\right)$$

Optimal Estimator $\delta: \mathcal{X} \to \mathbb{R}$ is said to be an estimator fn. $\psi: \Theta \to \mathbb{R}$ if it is statistic

MSE fn. $E_\theta\left[(\delta(\bar{x}) - \psi(\bar{\theta}))^2\right] = MSE_\delta(\bar{\theta})$

Claim: $MSE(\bar{\theta}) = Var_{\bar{\theta}}(\delta(\bar{x})) + \left(E_\theta[\delta(\bar{x})] - \psi(\bar{\theta})\right)^2$

if $\psi(\bar{\theta}) = E_\theta[\delta(\bar{x})] \ \forall \theta \in \Theta$, then '$\delta$' is called unbiased estimator

Lecture 27

Thursday, October 5, 2017 9:36 AM

$T_1, \dots, T_m$ are unbiased estimators,
then $\sum_k \alpha_k T_k$ is an unbiased estimator if $\sum_k \alpha_k = 1$

→ $X \sim$ Bernoulli (p)
$\psi(p) = p^2$
Unbiased estimator for ψ doesn't exist

→ MSE for an estimator T.
$$E_{\bar{\theta}}[(T(\bar{x}) - \psi(\bar{\theta}))^2] = var_{\bar{\theta}}(T(\bar{x})) + (E_{\bar{\theta}}[T(\bar{x})] - \psi(\bar{\theta}))^2$$

We want to find unbiased estimator which minimizes MSE

Let $\mathcal{U}$ denote a class of unbiased estimators for a given $\{F_{\bar{\theta}}; \bar{\theta} \in \Theta\}$
and a fn. ψ
Find $T^* \in \mathcal{U}$ st. $\forall T \in \mathcal{U}$, $var(T^*) \le var_{\bar{\theta}}(T)$ $\forall \bar{\theta} \in \Theta$

Ex $X_1, \dots, X_n \sim$ Bernoulli (p), iid R.V.
Let $\psi(p) = p$

$\Theta = [0, 1]$ → parameter space

$T = \sum_k \alpha_k X_k$; $\sum_k \alpha_k = 1$

$E_p[T(\bar{\alpha})] = p$

$$Var_p(T(\bar{\alpha})) = E_p[(T(\bar{\alpha}) - p)^2]$$
$$= E_p[(\sum \alpha_k (x_k - p))^2]$$
$$= \sum \alpha_k^2 E(x_k - p)^2 = \sum_k \alpha_k^2 p$$

∴ We have to minimize $\sum_k \alpha_k^2$ for $\sum_k \alpha_k = 1$

Claim that optimal sol^n is $\alpha_k = \frac{1}{n}$ $\forall k$

Let $\mathcal{U}$ denote a class of unbiased estimators for a given $\{F_{\bar{\theta}}; \bar{\theta} \in \Theta\}$
and a fn. ψ; & $\mathcal{U}_0$ be the same for $\{F_{\bar{\theta}}; \bar{\theta} \in \Theta\}$ & $\psi = 0$

$T \in \mathcal{U}$ is optimal iff $E_{\bar{\theta}}[Tv] = 0$ $\forall$ $v \in \mathcal{U}_0$, $\bar{\theta} \in \Theta$

10/9/2017 11:38 AM

Aim: Find $T^* \in \mathcal{U}$ st. $var_{\bar{\theta}}(T^*) \le Var_{\bar{\theta}}(T)$ $\forall \bar{\theta} \in \Theta$ & $T \in \mathcal{U}$

T^* is called Uniform Minimal Variance Unbiased Estimator (UMVUE)

Ex Bernoulli$(p) \sim X_1, \dots, X_n$
$T(\bar{x}) = X_k$ is unbiased for $\psi(p) = p$ $\forall k$
⇒ $\sum_k \alpha_k X_k$ is also unbiased for $\sum_k \alpha_k = 1$

Result. T^* is UMVUE iff $E_{\bar{\theta}}[T^* v] = 0$ $\forall$ $v \in \mathcal{U}_0$ & $\theta \in \Theta$

Obs... [illegible]

Result. T^* is UMVUE iff $E_{\bar\theta}[T^* v] = 0 \ \forall\ v \in \mathcal{U}_0$ & $\theta \in \Theta$

Observation: if $T \in \mathcal{U}$ & $v \in \mathcal{U}_0$, then $T + cv \in \mathcal{U}\ \forall\ c \in \mathbb{R}$

Proof: Part I: if T^* is UMVUE then $E_{\bar\theta}[T^* v] = 0\ \forall\ \bar\theta \in \Theta\ \forall\ v \in \mathcal{U}_0$
(By contradiction)
Suppose not. Then, $\exists\ v_0 \in \mathcal{U}_0$ & $\bar\theta \in \Theta$ s.t. $E_{\bar\theta}[T^* v] \neq 0$
Consider $T = T^* - cv_0 \in \mathcal{U}\ \forall c \in \mathbb{R}$

$$\mathrm{Var}_{\theta_0}(T) = \underbrace{E_{\theta_0}(T^2)}_{\Downarrow} - [\psi(\theta_0)]^2$$

$$E_{\theta_0}[(T^* - cv_0)^2] = E_{\theta_0}[(T^*)^2 - 2cv_0 T^* + c^2 v_0^2]$$

$$= \underbrace{E_{\theta_0}[(T^*)^2]}_{} - 2c\, E_{\theta_0}[v_0 T^*] + c^2 E_{\theta_0}[v_0^2]$$

$$= \mathrm{Var}(T^*) + [\psi(\theta_0)]^2$$

$$\Rightarrow \mathrm{Var}(T) = \mathrm{var}(T^*) - 2c\, E_{\theta_0}[T^* v_0] + c^2 E_{\theta}[v_0^2]$$

$$c = \frac{E_{\theta_0}[T^* v_0]}{E_{\theta}[v_0^2]} \quad \Rightarrow \mathrm{Var}(T) = \mathrm{var}(T^*) - \left(\frac{E^2_{\theta_0}[T^* v_0]}{E_{\theta_0}[v_0^2]}\right) > 0$$

$\Rightarrow$ $\mathrm{var}(T) < \mathrm{Var}(T^*)$ Contradiction to def. of T^*

$\therefore$ T^* is UMVUE $\Rightarrow E_{\bar\theta}[T^* v] = 0\ \forall\ v \in \mathcal{U}_0$ & $\bar\theta \in \Theta$

10/10/2017 8:42 AM

Part 2: Given: $T \in \mathcal{U}$; $E_\theta[Tv] = 0\ \forall\ \theta\ \forall\ v \in \mathcal{U}_0$
Suppose $\exists\ T_0 \in \mathcal{U}$; $v = T - T_0 \in \mathcal{U}_0$

$$E_\theta[T(T - T_0)] = 0 \Rightarrow E_\theta[T^2] = E_\theta[TT_0] \quad -(1)$$

$$\mathrm{var}_\theta(T) = E_\theta[T^2] - E_\theta^2[T]$$
$$= E_\theta[T^2] - \psi^2(\theta) \quad -(2)$$

$$\mathrm{Cov}_\theta(T, T_0) = E_\theta[(T - \psi(\theta))(T_0 - \psi(\theta))] = E_\theta[TT_0] - \psi^2(\theta) \quad -(3)$$

$$\therefore \mathrm{Cov}_\theta(T, T_0) = \mathrm{Var}_\theta(T) \quad [\text{from } (1), (2) \ \&\ (3)]$$

$$\rho_\theta(T, T_0) = \frac{\mathrm{Cov}_\theta(T, T_0)}{\sqrt{\mathrm{Var}_\theta(T)}\sqrt{\mathrm{Var}_\theta(T_0)}}$$

$$= \frac{\mathrm{var}_\theta(T)}{\sqrt{\mathrm{Var}_\theta(T)}\sqrt{\mathrm{var}_\theta(T_0)}} = \sqrt{\frac{\mathrm{Var}_\theta(T)}{\mathrm{var}_\theta(T_0)}} \leq 1 \quad (\because \rho_\theta \leq 1)$$

$\Rightarrow$ $\mathrm{var}_\theta(T) \leq \mathrm{var}_\theta(T_0) \Rightarrow T$ is UMVUE since $T_0 \in \mathcal{U}$

Ex Bernoulli(p); $p \in [0,1]$; $\psi(p) = p$

$T[\bar x] = \frac{1}{n}\sum_k x_k$. $\boxed{\text{Check if } T \text{ is UMVUE}; \forall\, p \in [0,1]\ \forall\ v \in \mathcal{U}_0\ E_p[Tv] = 0}$

$$E_p[Tv] = E_p[T(\bar x)\, v(\bar x)] = \sum_{\bar x \in \{0,1\}^n} T(\bar x)\, v(\bar x)\, P_p(\bar X = \bar x)$$

$$= \frac{1}{n}\sum_{\bar x}\left(\sum_k x_k\right) v(\bar x)\, p^{\sum x_i}(1-p)^{n - \sum x_i}$$

To prove: $\sum_{\bar x}\left(\sum_k x_k\right) v(\bar x)\, \alpha^{\sum x_i} = 0 \quad \forall\ \alpha \in (0, \infty)$

We know that $\upsilon \in \mathcal{U}_0$

$\therefore E_p[\upsilon] = 0 \quad \forall p \in (0,1)$

$$\Rightarrow \quad 0 = \sum_{\bar{x}} \upsilon(\bar{x})\, p^{\sum x_i} (1-p)^{n-\sum x_i}$$

$$\Rightarrow \quad 0 = \sum_{\bar{x}} \upsilon(\bar{x})\, \alpha^{\sum x_i} \quad ; \text{Diff. w.r.t. } \alpha$$

$$0 = \sum_{\bar{x}} \upsilon(\bar{x}) \times \sum x_i \times \alpha^{\sum x_i}$$

Hence proven

Lecture 30

Sunday, October 15, 2017 6:36 PM

- UMVUE recipe
  1) Sufficient statistic
  2) Unbiased estimate
  3) Rao-Blackwell Theorem
  4) Notion of "completeness
  5) Lehmann-Scheffe Theorem

Rao-Blackwell Theorem

If T is an unbiased estimator of $\psi(\theta)$ and S is sufficient for $\{f_\theta : \theta \in \Theta\}$
Then,
1) $E[T|S]$ is a unbiased statistic
2) $E_\theta[(E_\theta[T|S]-\psi(\theta))^2] \le E_\theta[(T-\psi(\theta))^2]$ i.e. $MSE_\theta(E[T|S]) \le MSE_\theta(T)$

Proof: 1) $E[T|S]$ is a statistic

Let $z = E[X|Y]$

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

$= \int_{-\infty}^{\infty} x f_{X|Y}(x|Y=y)\,dx$

$z = \int_{-\infty}^{\infty} x \frac{f_{XY}(x,y)}{f_Y(y)}\,dx = fn(y) \ne fn(\theta)$

$\therefore z$ is not a fn. of parameter θ

$\therefore E[T|S]$ is a statistic

2) $E[T|S]$ is an unbiased statistic

$E[E(X|Y)] = E[X]$ i.e. the Chain rule of expectation

$\therefore E[E[T|S]] = E[T] = \psi(\theta)$ $\because T$ is unbiased

$\therefore E[T|S]$ is an unbiased statistic

3) $MSE_\theta(E[T|S]) \le MSE_\theta(T)$

$z = E[T|S]$

$LHS = E[(z-Ez)^2] = E[z^2] - [Ez]^2 = E[z^2] - (\psi(\theta))^2$

$RHS = E[T^2] - (\psi(\theta))^2$

T.P. $\rightarrow E[z^2] \le E[T^2]$

$E[E[T^2|S]] = E[T^2]$ by chain rule of expectation

T.P. $\rightarrow E[(E[T|S])^2] \le E[E[T^2|S]]$

$(E[T|S])^2 \le E[T^2|S] \Rightarrow$ ↑

T.P. $\rightarrow (E[T|S])^2 \le E[T^2|S]$

$X \sim f_{T|S}(t|s)$

$E[T|S] = \int_{-\infty}^{\infty} t\, f_{T|S}(t|s)\,dt = E[X]$

$E[T^2|S] = \int_{-\infty}^{\infty} t^2 f_{T|S}(t|s)\,dt = E[X^2]$

We know that $(E[X])^2 \le E[X^2]$ (equal for var=0)

$\therefore (E[T|S])^2 \le E[T^2|S]$

Hence proven

Completeness

Statistic T is complete, if for any $g: \mathbb{R} \to \mathbb{R}$ such that

$$E_\theta[g(T)] = 0 \quad \forall \theta \in \Theta$$

then, $P_\theta[g(T) = 0] = 1 \quad \forall \theta \in \Theta$

Lecture 31

Saturday, October 14, 2017 9:36 AM

Recap

→ Blackwell Rao Theorem:

If T is an unbiased estimator of θ & S is a sufficient statistic for $\{f_\theta : \theta \in \Theta\}$

then, i) $E[T|S]$ is also an unbiased estimator

ii) $\text{var}_\theta(E[T|S]) \le \text{var}_\theta(T) \;\forall \theta \in \Theta$

→ Complete statistic

If a fn. g s.t. $E_\theta[g(T)] = 0 \;\forall \theta$, then $P_\theta[g(T) = 0] = 1 \;\forall \theta$

Ex $S = \max\{X_1, X_2, \ldots, X_n\}$; $X_i \sim \text{Uniform}(0,\theta)$; $F_{S,\theta}(x) = P_\theta(\max\{X_1, \ldots, X_n\} \le x)$

$= P_\theta(X_1 \le x, \ldots, X_n \le x) = \prod P_\theta(X_k \le x)$

$$F_{S,\theta}(x) = \begin{cases} 1 ; \text{ if } x \ge \theta \\ 0 ; \text{ if } x \le 0 \\ \left(\frac{x}{\theta}\right)^n ; \text{ if } 0 < x < \theta \end{cases}$$

$$f_{S,\theta}(x) = \begin{cases} \frac{n x^{n-1}}{\theta^n} ; \text{ if } x \in (0,\theta) \\ 0 ; \text{ if } x \notin (0,\theta) \end{cases}$$

Let g be a fn. s.t. $E_\theta[g(S)] = 0 \;\forall \theta$

$$\Rightarrow \frac{d}{d\theta}\int_0^\theta g(x)\, x^{n-1}\, dx = 0 \quad \forall \theta > 0$$

$$\Rightarrow \int_0^\theta g(x) \cdot \frac{n x^{n-1}}{\theta^n}\, dx = 0 \;\forall \theta > 0 \Rightarrow g(\theta)\theta^{n-1} = 0 \;\forall \theta > 0$$

⇒ S is complete sufficient statistic

Lehmann - Scheffè Theorem:

If S is complete sufficient & T is any unbiased estimator, then $E[T|S]$ is UMVUE

To prove: $\text{var}_\theta(E[T|S]) \le \text{var}_\theta(T') \;\forall T' \in \mathcal{U}\; \& \;\theta \in \Theta$

Proof: Let T_1 & $T_2 \in \mathcal{U}$, $E[T_1|S]$ & $E[T_2|S] \in \mathcal{U}$

$g(S) = E[T_1|S] - E[T_2|S] \in \mathcal{U}_0$

$E_\theta[g(S)] = 0 \;\forall \theta \Rightarrow$ 'S' is complete

$\Rightarrow P_\theta(g(S) = 0) = 1 \;\forall \theta$

$\therefore P(E[T_1|S] = E[T_2|S]) = 1 \;\forall \theta$

By Blackwell-Rao, $\text{var}_\theta(E[T_1|S]) \le \text{var}_\theta(T_1) \;\forall \theta$

and $E[T_1|S] = E[T|S] \;\forall T \in \mathcal{U}$

Ex $X_1, \ldots, X_n \sim \exp(\lambda)$; $\psi(\lambda) = \frac{1}{\lambda}$

$\sum X_k = S$ is sufficient ; $f_{T,\lambda}(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n-1)}$; Let g s.t. $E_\lambda[g(T)] = 0 \;\forall \lambda$

$\therefore \text{var}_\theta(E[T|S]) \le \text{var}(T') \;\forall T' \in \mathcal{U},\; \theta \in \Theta$

$$\int g(x) \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n-1)}\, dx = 0 \Rightarrow \int g(x)\, x^{n-1} e^{-\lambda x}\, dx = 0$$

$$L_h(\lambda) = \int^\infty h(x)\, e^{-\lambda x}\, dx \quad ; \quad L_h(\lambda) = 0$$

$\therefore h(x) \triangleq g(x)\, x^{n-1} = 0 \Rightarrow g(x) = 0 \;\forall x \ne 0$

Laplace transform is one-one i.e. Laplace = 0 ⇒ fn = 0

$L_h(\lambda) = \int_0^{\infty} h(x) e^{-\lambda x} dx$; $L_h(\lambda) = 0$

$\therefore h(x) \triangleq g(x) x^{n-1} = 0 \Rightarrow g(x) = 0 \ \forall x \neq 0$

Laplace ~~transform~~ is one-one, if laplace $= 0$, fn. $= 0$

$\therefore P(g(T)=0) = \cancel{1}$

$\therefore S$ is complete, $\therefore E[T|S]$ is UMVUE

$\Rightarrow \int_0$

$h(x)$

Cramer – Rao Variance Bound

likelyhood fn. $L(\theta, \bar{x}) = f_\theta(\bar{x})$

$X_1, X_2, \ldots, X_n$ iid. $\max_\theta \{L(\theta, \bar{x})\} = f_n(\bar{x})$

Let $x_1, \ldots, x_n$ be the observed values

$$\theta^* = \arg\max_{\theta \in \Theta} \{L(\theta, \bar{x})\}$$

θ^* is called Maximum likelyhood estimator

Ex Bernoulli (p) ; $L(p, \bar{x}) = p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$

$$\tilde{L} = \log p^{\Sigma x_k}(1-p)^{n-\Sigma x_k} = \Sigma x_k \log p + (n - \Sigma x_k)\log(1-p)$$

$$\frac{\partial}{\partial p}\tilde{L} = \Sigma x_k \frac{1}{p} + (n - \Sigma x_k)\frac{(-1)}{1-p} = 0 \Rightarrow (1-p)\Sigma x_k = p(n - \Sigma x_k) \Rightarrow p = \frac{1}{n}\Sigma x_k$$

Ex $X_1, X_2, \ldots, X_n \sim$ uniform $(0, \theta)$

MLE $= \max(x_k)$

$$L(\theta, \bar{x}) = \frac{1}{\theta^n} 1_{\min x_k \geq 0} 1_{\max x_k \leq \theta}$$

Lecture 33

Tuesday, October 17, 2017 8:42 AM

Recap

Hypothesis testing

- Parameter space Θ, sample space $\mathcal{X}$
- Null hypothesis (H_0) : $\theta \in \Theta_0$
- Alternate hypothesis (H_1) : $\theta \in \Theta_1 = \Theta_0^c$
- Test $\varphi : \mathcal{X} \to \{0,1\}$ or $\varphi : \mathcal{X} \to [0,1]$

— x —

for $\varphi : \mathcal{X} \to \{0,1\}$

$C_\varphi = \{\bar{x} : \varphi(\bar{x}) = 1\}$, $C_\varphi \subseteq \mathcal{X}$

Set of all samples for which the test chooses H_1

C_φ is called critical region of φ

φ \ Truth	H_0	H_1
H_0	✓	Type II error
H_1	Type I error	✓

for $\theta \in \Theta_0$, $E_\theta[\varphi(\bar{x})] = P_\theta(\varphi(\bar{x}) = 1)$

$= P_\theta(\bar{x} \in C_\varphi) =$ probability of Type I error

$= \beta_\theta(\varphi)$ (for $\theta \in \Theta_0$)

Level of Significance α:

Test φ is said to have significance α if

$\beta_\theta(\varphi) \le \alpha \;\; \forall \theta \in \Theta$

$\Phi(\alpha, \Theta_0, \Theta_1)$ is a class of tests with level of significance α

$P_\theta(\text{type II error}) = P_\theta(\varphi(\bar{x}) = 0)$, for $\theta \in \Theta_1$

$= P_\theta(\bar{x} \in C_\varphi^c)$

$= 1 - \beta_\theta(\varphi)$ (for $\theta \in \Theta_1$)

$\beta_\theta(\varphi) = E_\theta[\varphi(\bar{x})]$

$= P(\text{type I error})$ if $\theta \in \Theta_0$

$= 1 - P(\text{type II error})$ if $\theta \in \Theta_1$

$\beta_\theta(\varphi)$ is called the power of test φ

Lecture 34

Friday, October 27, 2017 1:21 PM

Hypothesis testing problem (α, H_0, H_1)

- test φ is feasible iff $\beta_\varphi(\bar\theta) \le \alpha \;\forall \bar\theta \in \Theta_0$
  $\varphi \in (\alpha, H_0, H_1)$

- test φ is most powerful for $\bar\theta \in \Theta_1$ if

$$\beta_\varphi(\bar\theta) \ge \beta_{\varphi'}(\bar\theta) \;\forall \varphi' \in (\alpha, H_0, H_1)$$

(i.e. P_θ(type II error for φ) $\le P_\theta$(type I error for φ')

- test $\varphi \in (\alpha, H_0, H_1)$ is uniformly most powerful if φ is most powerful $\forall \theta \in \Theta_1$ — May not exist in some cases

ex $X_1, X_2, \ldots, X_n \sim G(\mu, 1)$, $\mu \in \{a, -a\}$

$H_0 := \mu = a$
$H_1 := \mu = -a$
Simple Hypothesis

$$\varphi_\varepsilon \equiv \varphi(\bar x) = 0 \text{; if } \sum_n x_n \ge \varepsilon$$
$$= 1 \text{; else}$$

$$\beta_{\varphi_\varepsilon}(a) = E_a[\varphi_\varepsilon(\bar x)] = P_a(\varphi_\varepsilon(\bar x) = 1)$$

$$= P_a\Big(\underbrace{\frac{\Sigma x_k}{n}}_{Y \sim G(a, 1/n)} < \varepsilon\Big)$$

$$= P_a(Y < \varepsilon)$$

$$= P_a((-Y) > -\varepsilon) = P_a\Big(\underbrace{\frac{(-Y)-(-a)}{1/\sqrt n}}_{\sim G(0,1)} > \frac{-\varepsilon + a}{1/\sqrt n}\Big)$$

$$= ERFC(\sqrt n (a - \varepsilon))$$

Let $z_\alpha \in \mathbb{R}$ st $ERFC(z_\alpha) = \alpha$
for req. significance level, $z_\alpha \le \sqrt n(a-\varepsilon)$
$\Rightarrow \varepsilon \le a - \frac{z_\alpha}{\sqrt n}$

—x—

Uniformly most powerful test doesn't exist in general.
It exists in the foll. special cases:
H_0 & H_1 are both simple (i.e. have 1 element each)
(didn't do any more?)

Generalisation of test

$\varphi : \mathcal{X} \to \{0,1\}$
$\varphi : \mathcal{X} \to [0,1]$

Reject H_0 & H_1 with probability $\varphi(\bar x)$

Do not reject H_0 w/ probability $1-\varphi(\bar{x})$

→ $\bar{X} = \begin{cases} \bar{x}_1 \text{ ; w/ prob. } q \\ \bar{x}_2 \text{ ; w/ prob. } 1-q \end{cases}$

$\varphi(\bar{x}_1) = p_1$ and $\varphi(\bar{x}_2) = p_2$

Y is the output, so $Y \in \{0,1\}$

- For $\bar{x}_1$, $P(Y=1) = 1 - P(Y=0) = p_1$

for $\bar{x}_2$, " = " $= p_2$

- $\beta_\varphi(\bar{\theta}) = E_{\bar{\theta}}[\varphi(\bar{x})]$

$= \varphi(\bar{x}_1) P(\bar{X}=\bar{x}_1) + \varphi(\bar{x}_2) P(\bar{X}=\bar{x}_2)$

$= p_1 q + p_2(1-q)$

$\beta_\varphi(\bar{\theta}) = \int \varphi(\bar{x}) f_\theta(\bar{x})\, d\bar{x}$ Law of unconcious statistician

If $\bar{\theta} \in \Theta_0$, $P_{\bar{\theta}}(Y=1) = P_{\bar{\theta}}(Y=1 \mid \bar{X}=\bar{x}_1) P(\bar{X}=\bar{x}_1) + P_{\bar{\theta}}(Y=1 \mid \bar{X}=\bar{x}_2) P(\bar{X}=\bar{x}_2)$

$= p_1 q + p_2(1-q)$

$= \beta_\varphi(\bar{\theta})$

Even in this case, $\beta = P(\text{type II})$ if $\bar{\theta} \in \Theta_0$

$= 1 - P(\text{type I})$ if $\bar{\theta} \in \Theta_1$

Nayman-Pearson Lemma

Part A Any test φ of the form (for $\gamma : \mathcal{X} \to [0,1]$ & $k \geq 0$)

$$\varphi(\bar{x}) = \begin{cases} 1 \text{ ; if } f_{\theta_1}(\bar{x}) > k f_{\theta_0}(\bar{x}) \\ \gamma(\bar{x}) \text{ ; if } f_{\theta_1}(\bar{x}) = k f_{\theta_0}(\bar{x}) \\ 0 \text{ ; else} \end{cases}$$

is a maximum power test for its size (α)

- Significance level $\alpha_{k,\gamma} = E_\theta[\varphi(\bar{x})]$

$= P_{\theta_0}(f_{\bar{\theta}_1}(\bar{x}) > k f_{\bar{\theta}_0}(\bar{x})) + \int_{\bar{x} = f_{\theta_1}(\bar{x}) = k f_{\theta_0}(\bar{x})} \gamma(\bar{x}) f_{\theta_0}(\bar{x})\, d\bar{x}$

Part B For any given significance level α, $\exists$ k & γ st the test described is MPT, where γ is a constant

Lecture 36

Thursday, October 26, 2017 9:35 AM

Recap:

- Both H_0 & H_1 are simple
  find the MP test

  $H_0: \theta = \theta_0$; $X_1, X_2, \dots, X_n \sim f_{\theta_0}$

  $H_1: \theta = \theta_1$; $X_1, X_2, \dots, X_n \sim f_{\theta_1}$

Neyman-Pearson Lemma

(a) Consider a test of the foll. type for a const. $k \geq 0$ and a fn. $\gamma: \mathcal{X} \to [0,1]$.

$$\varphi(\bar{x}) = \begin{cases} 1 & ; \text{ if } f_{\theta_1}(\bar{x}) > k f_{\theta_0}(\bar{x}) \\ \gamma(\bar{x}) & ; \text{ if } f_{\theta_1}(\bar{x}) = k f_{\theta_0}(\bar{x}) \\ 0 & ; \text{ otherwise} \end{cases}$$

φ is MP for its size

(b) α is the req. size. Then the test ($k \geq 0, \gamma \in [0,1]$)

$$\varphi(\bar{x}) = \begin{cases} 1 & ; \text{ if } f_{\theta_1}(\bar{x}) > k f_{\theta_0}(\bar{x}) \\ \gamma & ; \text{ if } f_{\theta_1}(\bar{x}) = k f_{\theta_0}(\bar{x}) \\ 0 & ; \text{ else} \end{cases}$$

is MP for size α, when k & γ satisfy $E_{\theta_0}[\varphi(\bar{x})] = \alpha$

Ex X is a discrete r.v.

X	1	2	3	4	5	6
$f_0(\bar{x})$	0.01	0.01	0.01	0.01	0.01	0.95
$f_1(x)$	0.05	0.04	0.03	0.02	0.01	0.85

; Given dice throw, guess dice 1 or dice 2

Design a test w/ $\alpha = 0.03$

Test $\varphi(\bar{x}) = \begin{cases} 1; & \text{if } \bar{x} \in \{3,4,5\} \\ 0; & \text{else} \end{cases}$

$E_{\theta_0}[\varphi(x)] = 1 \times P_{\theta_0}(\varphi(x) = 1) + 0 \times P_{\theta_0}(\varphi(x) = 0)$

$= P_{\theta_0}(\varphi(x) = 1) = 0.03$

10/30/2017 11:41 AM

Ex $X_1, X_2, \dots, X_n \sim \text{Bernoulli}(p)$

$H_0: p = p_0 \quad H_1: p = p_1$

Given $\alpha \in [0,1]$

$$P_{p_0}(\bar{X} = \bar{x}) = p_0^{\Sigma x_i} (1-p_0)^{n - \Sigma x_i}$$

$$P_{p_1}(\bar{X} = \bar{x}) = p_1^{\Sigma x_i} (1-p_1)^{n - \Sigma x_i}$$

$$\lambda(\bar{x}) = \frac{P_{p_1}}{P_{p_0}} = \left(\frac{p_1}{p_0}\right)^{\Sigma x_i} \left(\frac{1-p_1}{1-p_0}\right)^{n - \Sigma x_i}$$

$\lambda(\bar{x})$ is a monotone fn. of Σx_i, inc. for $p_1 > p_0$

Eg $X_1, X_2, \ldots, X_n \sim G(\mu, 1)$

$H_0: \mu = \mu_0$, $H_1: \mu = \mu_1$ $(\mu_0 < \mu_1)$

$$\lambda(\bar{x}) = \frac{f_{\mu_1}(\bar{x})}{f_{\mu_0}(\bar{x})} = \prod_{i=1}^{n} \frac{f_{\mu_1}(x_i)}{f_{\mu_0}(x_i)} = e^{-n(\mu_1^2 - \mu_0^2)} \, e^{2(\mu_1 - \mu_0)\sum x_i}$$

$\lambda(\bar{x})$ is a monotone inc. fn. of $\sum x_i$

Test $\varphi = \begin{cases} 1; & \text{if } \sum x_i \geq z \\ 0; & \text{otherwise} \end{cases}$

$$E_{\mu_0}[\varphi(\bar{x})] = P_{\mu_0}(\varphi(\bar{x}) = 1) = P_{\mu_0}\left(\sum x_i \geq z\right)$$

$$= P_{\mu_0}\left(\frac{\sum x_i - n\mu_0}{n} \geq \frac{z - n\mu_0}{n}\right) = Erfc\left(\frac{z - n\mu_0}{n}\right)$$

$$Erfc\left(\frac{z - n\mu_0}{n}\right) \leq \alpha$$

31/10/17 - Tuesday Lec 38

Interval Estimation

find $L: \mathcal{X} \to \mathbb{R}$ & $U: \mathcal{X} \to \mathbb{R}$

interval $S(\bar{x}) = [L, U]$

"Good" event — $\theta \in S(\bar{x})$

"Bad" event — $\theta \notin S(\bar{x})$

Find $S(\bar{x}) = [L(\bar{x}), U(\bar{x})]$ s.t. $P_\theta(\theta \in S(\bar{x})) \geq 1 - \alpha \quad \forall \theta \in \Theta$

lower confidence ; upper confidence
bound at level
of confidence $(1-\alpha)$

Lecture 39

Thursday, November 2, 2017 9:42 AM

Confidence intervals

$X_1, X_2, \ldots, X_n \sim f_\theta(\cdot)$ iids , $\theta \in \mathbb{R}$

find $L(\bar{x})$ & $U(\bar{x})$, equivicantly

$$S(\bar{x}) = [L(\bar{x}), U(\bar{x})] \text{ st } P_\theta(\theta \in S(\bar{x})) \geq 1-\alpha \ \forall \theta$$

and given a value of α

ex $X_1, X_2, \ldots, X_n \sim G(\mu, 1)$

$U(\bar{x}) = \frac{1}{n}\sum x_i + c$;

$L(\bar{x}) = \frac{1}{n}\sum x_i - c$; $c > 0$

By choosing appropriate c, we can get any level of confidence 'α'

Uniformly Most Accurate

The collection $S(\bar{x})$ is said to be UMA if

$$P_\theta(\theta' \in S(\bar{x})) \leq P_\theta(\theta' \in S'(\bar{x})) \ \forall \ \theta' \neq \theta \text{ \& interval collection } S'(\bar{x})$$

Pivot

$T : \mathcal{X} \times \Theta \to \mathbb{R}$

any such fn. is called pivot if distribution of $T \perp \theta$

ex $X_1, X_2, \ldots, X_n \sim G(\mu, 1)$

$T(\bar{x}, \mu) = \frac{1}{n}\sum x_i - \mu \sim G(0, 1/n)$

Find c_1 & c_2 st. $P_\theta(c_1 \leq T(\bar{x}, \mu) \leq c_2) \geq 1-\alpha$

$(\cdot) = P_\mu(c_1 \leq \frac{1}{n}\sum x_i - \mu \leq c_2)$

$= P_\mu(c_1 + \mu \leq \frac{1}{n}\sum x_i \leq c_2 + \mu)$

6/11 – Lect 40

Solve $P(-c \leq \bar{x} - \mu \leq c) = P(\bar{x} - c \leq \mu \leq \bar{x} + c)$

LHS $\neq$ fn.(μ), solve for c using LHS $= 1-\alpha$

ex $X_1, X_2, \ldots, X_n \sim \text{Uniform}(0, \theta)$

find α-confidence interval for θ

$E[x_i] = \frac{1}{2}\theta$; $T(\bar{x}, \theta) = \frac{\sum x_i}{n} - \frac{\theta}{2}$

T isn't useful $\because$ c we get is a fn. of θ (T isn't a pivot)

Pivots are related to suff. statistics (usually)

$$T(x, \theta) = \frac{\max\{x_i\}}{\theta} = \max\left\{\frac{x_i}{\theta}\right\}$$

$$P\left(\max\left\{\frac{x_i}{\theta}\right\} \leq x\right) = \begin{cases} 0 ; \text{ if } x \leq 0 \\ x^n ; \text{ if } 0 < x < 1 \\ 1 ; \text{ if } x \geq 1 \end{cases}$$

$1 \; ; \text{ if } x = 1$

$$f_T(x) = \begin{cases} 0 \; ; \text{ if } x \notin [0,1] \\ nx^{n-1} \; ; \text{ if } x \in [0,1] \end{cases}$$

$$P(c \le T(\bar{x}, \theta) \le 1) = \int_c^1 nx^{n-1}\,dx = 1 - c^n$$

$\alpha = c^n$ for α-confidence level

$\Rightarrow c = \alpha^{1/n}$

$$(\alpha^{1/n} \le \overset{\max\{x_i\}}{y}/\theta \le 1) \equiv \left(\frac{1}{\alpha^{1/n}} \ge \frac{\theta}{y} \ge 1\right) \equiv \left(\frac{y}{\alpha^{1/n}} \ge \theta \ge y\right)$$

$\therefore$ Confidence interval for $\theta \equiv [y, y\alpha^{-1/n}]$

Inequalities & Central Limit Theorem

Method 1 $\rightarrow$ Chebyshev: $P(|X - E_x| > \varepsilon) \le \dfrac{var(X)}{\varepsilon^2}$

ex $X_1, X_2, \dots, X_n \sim$ Bernoulli(p) ; $p \in [0,1]$

$$P\left(\left|\frac{\Sigma X_i}{n} - p\right| > \varepsilon\right) \le \frac{var(Y)}{\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2} \le \frac{1}{4n\varepsilon^2} \neq fn.(p)$$

$$P\left(\frac{\Sigma X_i}{n} - c \le p \le \frac{\Sigma X_i}{n} + c\right) \ge 1-\alpha$$

$$P\left(\left|\frac{\Sigma X_i}{n} - p\right| \le c\right) \ge 1-\alpha$$

$$1 - P\left(\left|\frac{\Sigma X_i}{n} - p\right| > c\right) \ge 1-\alpha$$

$$P\left(\left|\frac{\Sigma X_i}{n} - p\right| > c\right) \le \alpha$$

$$\alpha = \frac{1}{4nc^2} \Rightarrow c = \sqrt{\frac{1}{4n\alpha}}$$

$\therefore$ Confidence Interval $\equiv \left[\dfrac{\Sigma X_i}{n} - \dfrac{1}{2\sqrt{n\alpha}}, \dfrac{\Sigma X_i}{n} + \dfrac{1}{2\sqrt{n\alpha}}\right]$

Method 2 $\rightarrow$ Central Limit Theorem

7/11 - Lect 41

Regression

- Establishes a relation b/w input & Output

Linear Regression

$$Y = \beta_0 + \sum_{i=1}^{r} \beta_i x_i$$

Aim $\rightarrow$ Estimate β_i

For indep. x_i, take $(r+1)$ obs and solve $(r+1)$ eqn.
This isn't deterministic & allows expt. error

$$Y = \beta_0 + \sum_{i=1}^{r} \beta_i x_i + E \; ; \; E \cdot \text{Error r.v.}, \text{ expectation} = 0$$

Simple linear regression

$$Y_i = \alpha + \beta x_i + \varepsilon_i$$

given: inputs (x_i) & responses (y_i)

Aim: Estimate α & β, s.t.

$$\text{Error} = \mathcal{E}(A,B) = \sum_{i=1}^{n} \left(Y_i - (A + Bx_i)\right)^2$$

A = estimate for α

B = estimate for β

to minimise $\mathcal{E}$, $\frac{\partial \mathcal{E}}{\partial A} = 0 \Rightarrow -2\sum (Y_i - (A + Bx_i)) = 0$

$$\frac{\partial \mathcal{E}}{\partial B} = 0 \Rightarrow -2\sum x_i (Y - (A + Bx_i)) = 0$$

$$A = \bar{y} - B\bar{x} \quad \& \quad B = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$