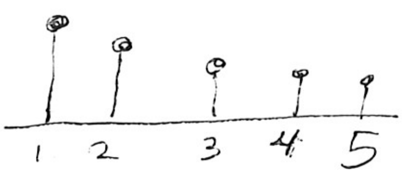
STAT 24400
Lecture 9
2/2/16

Review (for Midterm)

X a random variable

Discrete Case

$p(x) = P(X = x)$

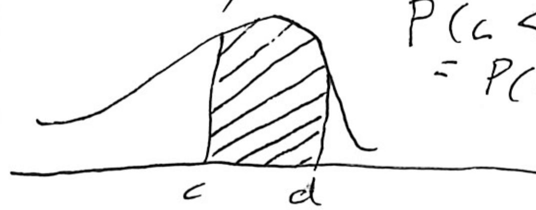


Ex Binomial
$b(x; n, \theta)$
(aka $Bin(x; n, \theta)$)
Negative Binomial
Bernoulli
Geometric

Continuous Case

density $f(x)$

$P(c < X \leq d)$
$= P(c \leq X \leq d)$



Ex: Uniform
Exponential
Beta
Normal
Poisson

cdf $F(x) = P(X \leq x)$

Transformations

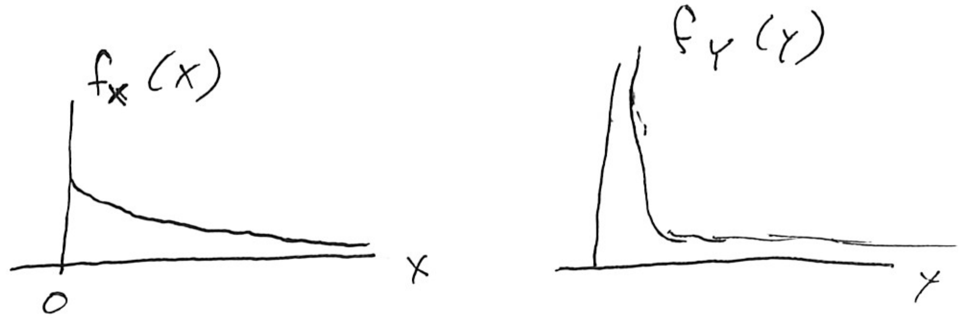
Given distribution of X ($p_X(x)$ or $f_X(x)$)

Find distribution of $Y = h(X)$ ($p_Y(y)$ or $f_Y(y)$)

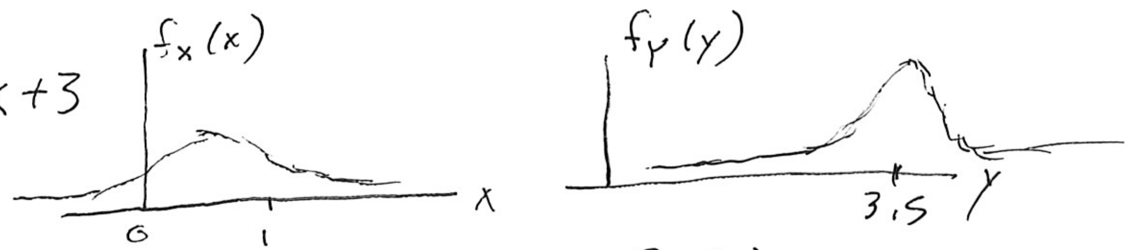
Ex: $Y = X^2$ $\quad Y = \log X$ $\quad$ etc

Transformation of Random Vars, continued - $Y = h(x)$

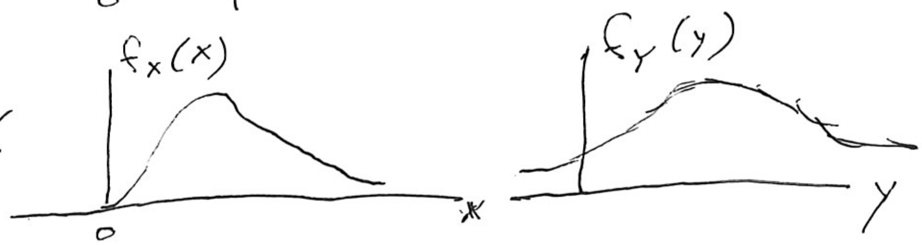
$Y = X^2$



$Y = \frac{1}{2}X + 3$



$Y = \ln X$



$Y = h(X) \qquad X = g(Y)$

$$f_Y(y) = f_X(g(y)) \cdot |g'(y)|$$

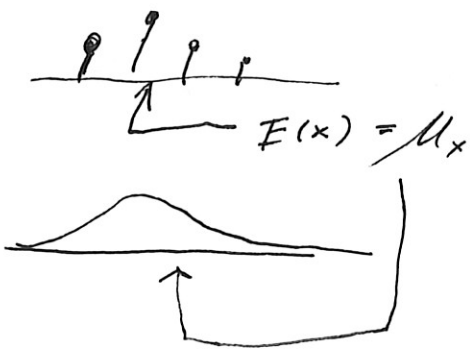
Discrete case: $p_X(x) = P_X(X = x)$

$p_Y(y) = p_X(g(y))$

Expectations

— "Center of Gravity"

$$E(X) = \begin{cases} \sum_{\text{all } x} x\, p(x) \\ \int_{-\infty}^{\infty} x\, f_X(x)\, dx \end{cases}$$



$$E(h(X)) = \begin{cases} \sum_{\text{all } x} h(x)\, p_x(x) \\ \int_{-\infty}^{\infty} h(x) f(x)\, dx \end{cases}$$

$E(h(X)) \neq h(E(x))$
UNLESS h linear
or X constant

Variance — "spread", "dispersion"

$$Var(X) = E\left((X-\mu_x)^2\right) = E(X^2) - (E(x))^2$$

Standard Deviation $X = \sqrt{Var(x)} = \sigma_x$

$$E(aX+b) = aE(X) + b$$

$[\text{or } \mu_{aX+b} = a\mu_x + b]$

$$Var(aX+b) = a^2\, Var(x)$$

$[\text{or } \sigma^2_{aX+b} = a^2 \sigma_x^2]$

Special Case: $a = \frac{1}{\sigma_x}$, $b = \frac{-\mu_x}{\sigma_x}$

Then:

$aX + b = \frac{x - \mu_x}{\sigma}$ "Standardized" Form

St'dized form (cont)

So:
$E(aX+b) = 0$
$Var(aX+b) = 1$

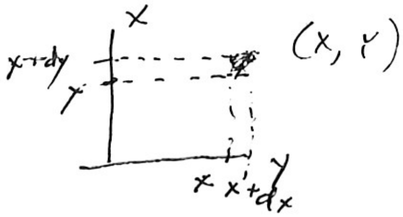
(3)

Multivariate Distributions

Bivariate

Discrete: $p(x, y) = P(X = x \text{ and } Y = y)$

Continuous: $f(x, y)\,dx\,dy = P(x < X \le x + dx,\ y \le Y \le y + dy)$



Marginal Distributions ("side view")

$$p_Y(y) = \sum_{\text{all } x} p(x, y), \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx$$

Conditional Distributions ("cross sections")

$$p(y|x) = \frac{p(x, y)}{p_x(x)}, \quad f(y|x) = \frac{f(x, y)}{f_x(x)}$$

Indep Random Vars

$p(x, y) = p_x(x)\,p_Y(y)$ for all x, y $\qquad$ $f(x, y) = f(x)\,f(y)$ for all x, y

$p_x(y) = p(y|x)$ for all x, y $\qquad$ $f_Y(y) = f(y|x)$ for all x, y

---

Note that Bivariate distributions determine marginal dists, but not other way around unless X and Y are independent

but $p_x(x)$ and $p(y|x)$ do determine $p(x, y)$
$f_x(x)$ and $f(y|x)$ " " by multiplication

Moment Generating Functions

The moment generating function (mgf) $M(t)$ is $E(e^{tx})$.

discrete

$$M(t) = \sum_{\text{all } x} e^{tx} p(x)$$

continuous

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

The r^{th} moment of a random var is $E(X^r)$.

$$M^{(r)}(0) = E(X^r)$$

the r^{th} derivative.

Trick to find moments by differentiating instead of integrating

$$\text{mgf} \rightleftarrows \text{pdf} \rightleftarrows \text{cdf}$$

Any one of the above gives the other two.

Bayes's Theorem

E_i : "causes" "states of nature"

F : "effect", "data"

$$P(E_i|F) = \frac{P(E_i)\,P(F|E_i)}{P(F)}$$

$$P(F) = \sum_{j=1}^{n} P(E_j)\,P(F|E_j)$$

Equivalently:

$$\underbrace{P(E_i|F)}_{\text{posterior}} \propto \underbrace{P(E_i)}_{\text{prior}}\,\underbrace{P(F|E_i)}_{\text{liklihood}}$$

(or $P(E_i|F) = (\text{Const.})\,P(E_i)\,P(F|E_i)$

$$\text{"Const"} = \frac{1}{P(F)} = \frac{1}{\sum_j P(E_j)\,P(F|E_j)}$$

Bayes's Theorem (continued)

To be clear, let's write in densities (or pmf's)

→ Given: $f_Y(y)$ and $f(x|y)$

→ Find: $f(y|x)$

These are the same !!

$$f(y|x) \propto f(x|y)\, f_Y(y)$$

$$f(y|x) = K\, f(x|y)\, f_Y(y) \qquad (K \text{ may depend on } x)$$

$$f(y|x) = \frac{f(x|y)\, f_Y(y)}{\int_{-\infty}^{\infty} f(x|u)\, f_Y(u)\, du}$$

Idea: Given X (data), make inferences about y

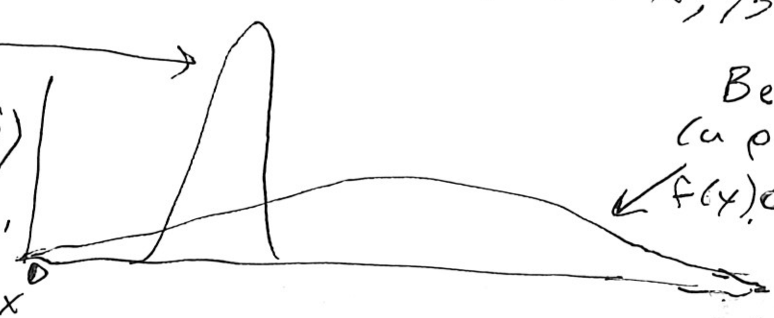
Ex

$f(y)$	Beta (α, β)
$p(x|y)$	Binomial (n, y)
$f(y|x)$	Beta $(\alpha + x, \beta + n - x)$

"After" (a posteriori)

$f(y|x) \propto y^{\alpha+x-1}(1-y)^{\beta+n-x-1}$

$E(Y) = \dfrac{\alpha + x}{\alpha+\beta+n}$



Before (a priori)

$f(y) \propto y^{\alpha-1}(1-y)^{\beta-1}$

$E(Y|X) = \dfrac{\alpha}{\alpha+\beta}$

Bayes for Binomial

θ = fraction, probability

$f(\theta)$ prior (Example: Beta(α, β))

$$f(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}, \quad 0<\theta<1$$

[$\alpha = \beta = 1$ gives uniform]

$$\text{Expectation} = \frac{\alpha}{\alpha+\beta} = \mu_\theta$$

$$\text{Variance} = \frac{\mu_\theta(1-\mu_\theta)}{\alpha+\beta+1}$$

$p(x/\theta)$ likelihood

$$p(x/\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x} \quad x = 0, 1, \dots n$$

$f(\theta/x)$ posterior

$$f(\theta/x) \propto \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}$$

Beta$(x+\alpha,\ n-x+\beta)$

$$E(\theta/X=x) = \frac{\alpha+x}{\alpha+\beta+n}$$

$$Var(\theta/X=x) = \frac{E(\theta/x)(1-E(\theta/x))}{\alpha+\beta+n+1}$$

Bayes for Normal

$$f(y/x) \propto f_Y(y)\, f(x/y)$$

$$\propto e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \cdot e^{-\frac{1}{2}(x-y)^2}$$

$$\Rightarrow e^{-\frac{1}{2}\frac{(y-A)^2}{B}}$$

by "completing the square"

$$ax^2 + bx + c \Rightarrow a(x+h)^2 + k$$

this gives

$$A = \frac{x\sigma^2}{\sigma^2+1} + \frac{\mu \cdot 1}{\sigma^2+1}$$ weighted average

$$B = \frac{\sigma^2}{\sigma^2+1}$$

Max Liklihood

Point estimation of $\hat{\theta}$, an estimate of θ

$$\text{Bias} = E(\hat{\theta}) - \theta$$

Mean squared Error $= MSE$

$$= E[(\hat{\theta} - \theta)^2]$$

$$= Var(\hat{\theta}) + (\text{Bias})^2$$

To find MLE:

Write down $L(\theta)$ (see lecture 8)

Take $\log L(\theta)$

want max $\log L(\theta)$:

set $\frac{d}{d\theta} \log L(\theta) = 0$, solve.

Check: find $\frac{d^2}{d\theta^2} \log L(\theta)$, check that

$$\frac{d^2}{d\theta^2} \log L(\theta) < 0 .$$