MS Theory-I

Lecture 5

Review:

Chapter 2: Transformations and Expectations

§ 2.1 Dist'ns of functions of a Random Variable (RV)
§ 2.2 Expected Values $E[X]$
§ 2.3 Moments and moment generating functions (mgf)
§ 2.4 Differentiating under an integral sign

§2.1 Transformations and Expectations

Goal: Know dist'n of X w/ sample space $\mathcal{X}$
Find dist'n of $Y = g(X)$ $\mathcal{Y}$

$$g(x): \mathcal{X} \rightarrow \mathcal{Y}$$
$$\leftarrow \; : g^{-1}(y)$$

Example 2.1.1 (Binomial Transform)

pmf: $f_X(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, 2 \ldots n$

$= \binom{n}{x} p^x (1-p)^{n-x} I_{[0,1,\ldots,n]}(x)$ $\begin{cases} n \text{ positive integer} \\ 0 \le p \le 1 \end{cases}$

indicator ft'n parameters

know, n,p
know X behavior

$Y = g(X)$, where $g(x) = n - x$ $\mathcal{X} = \{0, 1, 2, \ldots, n\}$
$y = n - x$ $\mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\}$
$x = n - y$ $= \{0, 1, \ldots, n\}$

for each x there is only one y.
each y at most one x.

$$\begin{pmatrix} x & y \\ 0 & n \\ 1 & n-1 \\ \vdots & \\ n & 0 \end{pmatrix}$$

← one to one and onto

$$f_Y(y) = \sum_{x \in g^{-1}(y)} f_X(x) = f_X(n-y) = \binom{n}{n-y} p^{n-y} (1-p)^{n-(n-y)}$$
$$= \binom{n}{y} (1-p)^y p^{n-y} * I_{[0,1,\ldots,n]}(y)$$

$$f_Y(y) = \binom{n}{y} (1-p)^y p^{n-y} * I_{[0,1,\ldots,n]}(y)$$

$$Y \sim \text{binomial}(n, (1-p))$$

Review cont: Continuous Distributions $<$ pdf, cdf

If transformation is from X to $Y = g(X)$

$$\mathcal{X} = \{x : f_X(x) > 0\} \quad \text{and} \quad \mathcal{Y} = \{y : y = g(x) \text{ for some } x \in \mathcal{X}\}.$$

pdf of random variable X is >0 only on $\mathcal{X}$ and zero else.
$\uparrow$ support set or support

Easiest deal with ft'ns $g(x)$ that are monotone $<$ increasing, decreasing

$x \to g(x)$ is monotone $\to$ one to one and onto

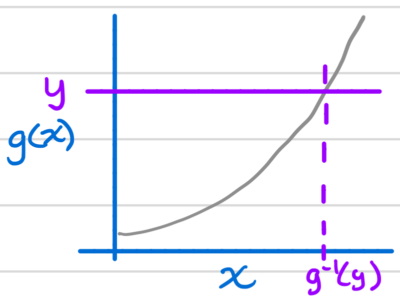
from $\mathcal{X} \to \mathcal{Y}$

each $x \longrightarrow$ one y
at most one x $\longleftarrow$ each y
$\underbrace{\qquad\qquad}_{\text{one-to-one}}$

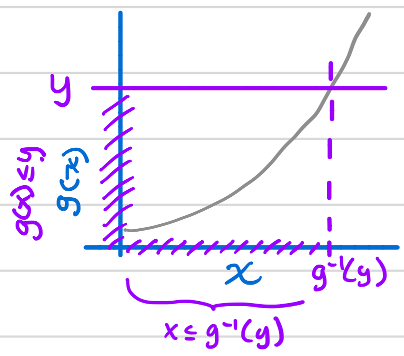
for each $y \in \mathcal{Y}$ there is an $x \in \mathcal{X}$ s.t. $g(x) = y$
$\underbrace{\qquad\qquad}_{\text{onto}}$

uniquely pairs x & y.

g(x) is increasing Find $F_Y = P(Y \le y)$

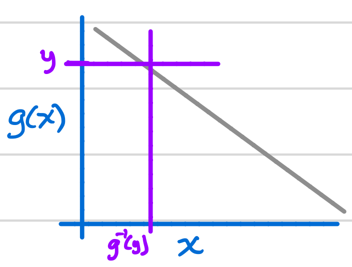


$$\{x \in \mathcal{X} : g(x) \le y\}$$
$$= \{x \in \mathcal{X} : g^{-1}(g(x)) \le g^{-1}(y)\}$$
$$= \{x \in \mathcal{X} : x \le g^{-1}(y)\}$$

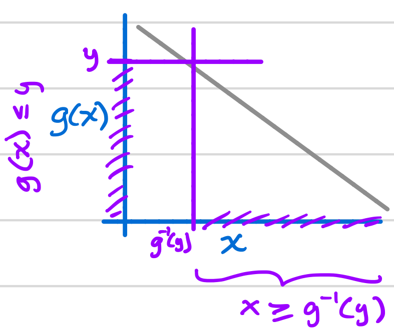


$$F_Y(y) = \int_{\{x \in \mathcal{X} : x \le g^{-1}(y)\}} f_X(x)\,dx = \int_{-\infty}^{g^{-1}(y)} f_X(x)\,dx = F_X(g^{-1}(y)).$$

g(x) is decreasing Find $F_Y = P(Y \le y)$



$$\{x \in \mathcal{X} : g(x) \le y\}$$
$$= \{x \in \mathcal{X} : g^{-1}(g(x)) \ge g^{-1}(y)\}$$
$$= \{x \in \mathcal{X} : x \ge g^{-1}(y)\}$$



$$F_Y(y) = \int_{g^{-1}(y)}^{\infty} f_X(x)\,dx = 1 - F_X(g^{-1}(y))$$

Review cont.

Theorem 2.1.3 *Let X have cdf $F_X(x)$, let $Y = g(X)$, and let $\mathcal{X}$ and $\mathcal{Y}$ be defined as in (2.1.7).*

a. *If g is an increasing function on $\mathcal{X}$, $F_Y(y) = F_X\left(g^{-1}(y)\right)$ for $y \in \mathcal{Y}$.*

b. *If g is a decreasing function on $\mathcal{X}$ and X is a continuous random variable, $F_Y(y) = 1 - F_X\left(g^{-1}(y)\right)$ for $y \in \mathcal{Y}$.*

2.1.7 : $\mathcal{X} = \{x : f_X(x) > 0\}$ and $\mathcal{Y} = \{y : y = g(x) \text{ for some } x \in \mathcal{X}\}$

Two new dist'ns :

Uniform Distribution (continuous)

Uniform(a, b)

pdf	$f(x\|a,b) = \frac{1}{b-a}$, $\underbrace{\overbrace{a}^{\text{parameters}} \le x \le \overbrace{b}}_{\text{Sample Space}}$
mean and variance	$\mathrm{E}X = \frac{b+a}{2}$, $\quad \mathrm{Var}\,X = \frac{(b-a)^2}{12}$
mgf	$M_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}$
notes	If $a = 0$ and $b = 1$, this is a special case of the beta ($\alpha = \beta = 1$).

Uniform (0,1) $\quad a = 0, b = 1$

$f(x \mid a=0, b=1) = \frac{1}{1-0} = 1 \quad 0 \le x \le 1$

$= 1 * I_{[0,1]}^{(x)}$

where $I_{[0,1]}^{(x)} = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{else} \end{cases}$

$X \sim$ Uniform (0,1) $\quad P(X \le x) = F_X(x) = \int_0^x 1\,dt = t\Big|_0^x = x$

Uniform $(0, \theta)$

pdf: $f_X(x \mid \theta) = \frac{1}{\theta} \quad 0 \le x \le \theta \quad = \frac{1}{\theta} I_{[0,\theta]}^{(x)}$ where $I_{[0,\theta]}^{(x)} = \begin{cases} 1 & 0 \le x \le \theta \\ 0 & \text{else} \end{cases}$

red flag sample space ft'n of parameter

Exponential distribution (continuous)

Exponential(β)

pdf	$f(x\|\beta) = \frac{1}{\beta} e^{-x/\beta}$, $\quad \overbrace{0 \le x < \infty}^{\text{sample space}}, \quad \overbrace{\beta > 0}^{\text{Parameter space}}$
mean and variance	$\mathrm{E}X = \beta$, $\quad \mathrm{Var}\,X = \beta^2$
mgf	$M_X(t) = \frac{1}{1-\beta t}$, $\quad t < \frac{1}{\beta}$
notes	Special case of the gamma distribution. Has the *memoryless* property. Has many special cases: $Y = X^{1/\gamma}$ is *Weibull*, $Y = \sqrt{2X/\beta}$ is *Rayleigh*, $Y = \alpha - \gamma \log(X/\beta)$ is *Gumbel*.

$X \sim$ exponential (β)

$$\underset{\text{cdf}}{F_X(x)} = P(X \le x) = \int_0^x \frac{1}{\beta} e^{-t/\beta}\,dt = -\frac{\beta}{\beta} e^{-t/\beta}\Big|_0^x = -e^{-x/\beta} + 1$$

$$= 1 - e^{-x/\beta}$$

Review cont.

Continuous Example (Uniform - exponential) — also one-to-one and onto.

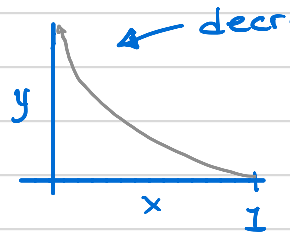
$X \sim f_X(x) = 1 \quad 0<x<1 = U(0,1)$ dist'n. $\qquad f_X(x) = 1 * I_{[0,1]}(x)$

$F_X(x) = \int_0^x 1\,dt = t\big|_0^x = x$

Interested in dist'n of $Y = g(X) = -\log_e(X)$

Sample space of X: $\mathcal{X} = \{x: 0 \le x \le 1\}$ $\qquad -\log(1) = 0; \ \lim_{B \to 0} (-\log(B)) = \infty$

$\rightarrow$ Y: $\mathcal{Y} = \{y: 0 \le y < \infty\}$



$y = -\log(x) = g(x)$

$g^{-1}(y) = e^{-y}$ (solve for x in terms of y).

$F_Y(y) = 1 - F_X(g^{-1}(y)) = 1 - F_X(e^{-y})$

$= 1 - e^{-y}$ (since $F_X = x$)

$F_Y(y) = 1 - e^{-y} \quad y > 0.$

$\hookrightarrow$ cdf of exponential $(B=1)$.

If pdf of Y is continuous obtain by differentiating cdf

continuous!

Theorem 2.1.5 *Let X have pdf $f_X(x)$ and let $Y = g(X)$, where g is a monotone function. Let $\mathcal{X}$ and $\mathcal{Y}$ be defined by (2.1.7). Suppose that $f_X(x)$ is continuous on $\mathcal{X}$ and that $g^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$. Then the pdf of Y is given by*

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \dfrac{d}{dy} g^{-1}(y) \right| & y \in \mathcal{Y} \\ 0 & \textit{otherwise.} \end{cases} \qquad (2.1.10)$$

Proof: (chain rule)

$$\underset{\text{pdf}}{f_Y(y)} = \frac{d}{dy} F_Y(y) = \begin{cases} f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) & g \text{ increasing} \\ -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) & g \text{ decreasing.} \end{cases}$$

Goal: Find dist'n of $Y = g(X)$, X, Y continuous

Approach:

Thm 2.1.5

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| & y \in \mathcal{Y} \\ 0 & \text{else} \end{cases}$$

Steps (continuous dist'n) for finding dist'n of $Y = g(X)$

- Determine if $g(x)$ is monotonic - Can we apply Thm 2.1.5?
- Find $g^{-1}(y)$: solve for x in terms of y.
- Determine Sample Space: $\mathcal{Y}$
- Calculate $\frac{d}{dy} g^{-1}(y)$ [derivative old wrt new]
- in $f_X(x)$ replace x with $g^{-1}(y)$ & multiply by $\left| \frac{d}{dy} g^{-1}(y) \right|$, identify $\mathcal{Y}$

$\Rightarrow$ pdf of Y: $f_Y(y)$

Example 2.1.6 (Inverted gamma pdf):

Gamma distribution (continuous)

Gamma(α, β)

pdf	$f(x\|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$,	$0 \le x < \infty$, $\alpha, \beta > 0$
mean and variance	$EX = \alpha\beta$, $\quad \text{Var}\, X = \alpha\beta^2$	
mgf	$M_X(t) = \left(\frac{1}{1-\beta t}\right)^\alpha$, $\quad t < \frac{1}{\beta}$	
notes	Some special cases are exponential ($\alpha = 1$) and chi squared ($\alpha = p/2$, $\beta = 2$). If $\alpha = \frac{3}{2}$, $Y = \sqrt{X/\beta}$ is *Maxwell*. $Y = 1/X$ has the *inverted gamma distribution*. Can also be related to the Poisson (Example 3.2.1).	

sample space

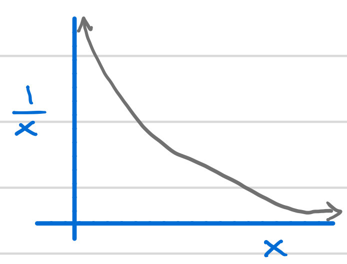
parameter space

$\Gamma(\alpha+1) = \alpha\Gamma(\alpha) \quad \alpha > 0$

$\Gamma(n) = (n-1)!$ if $n = 1, 2, 3 \ldots$

Let $f_X(x)$ be gamma(n, β): $f_X(x) = \frac{1}{(n-1)!\,\beta^n} x^{n-1} e^{-x/\beta} I_{(0,\infty)}(x)$ where $I_{(0,\infty)}(x) = \begin{cases} 1 & x \ge 0 \\ 0 & \text{else} \end{cases}$

Goal: Find dist'n of $g(X) = 1/X = Y$



- $g(x) = 1/x = y \Rightarrow x = 1/y = g^{-1}(y)$
- $\frac{d}{dx} g(x) = -1/x^2 < 0 \Rightarrow$ decreasing / monotonic

$$f_Y(y) = \frac{1}{(n-1)!\,\beta^n} \left(\frac{1}{y}\right)^{n-1} e^{-1/y\beta} * \left| \frac{-1}{y^2} \right| * I_{(0,\infty)}(y)$$

$\leftarrow$ sample space of y same as x.
as $x \to 0$ $y \to \infty$
$x \to \infty$ $y \to 0$

$$f_Y(y) = \frac{1}{(n-1)!\,\beta^n} \left(\frac{1}{y}\right)^{n+1} e^{-1/(\beta y)} \cdot I_{(0,\infty)}(y) \quad //.$$

$Y \sim$ inverted gamma

What if $Y = g(X)$ is not monotonic ft'n?

Example 2.1.7) (Square transformation)

- X is 'a continuous' Random Variable
- Assume $y > 0$ and we wish to determine the dist'n (cdf and pdf) of $Y = X^2$.

- $F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y})$ (since continuous)
$= P(-\sqrt{y} < X \le \sqrt{y})$
$= P(X \le \sqrt{y}) - P(X \le -\sqrt{y})$
$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$

Aside

Assume
$X \sim Normal(0,1)$ mean variance
pdf →

$P(X^2 \le 1) = P(-\sqrt{1} \le X \le \sqrt{1})$
$= P(-1 < X \le 1)$
$= P(X \le 1) - P(X \le -1)$
$= F_X(1) - F_X(-1)$

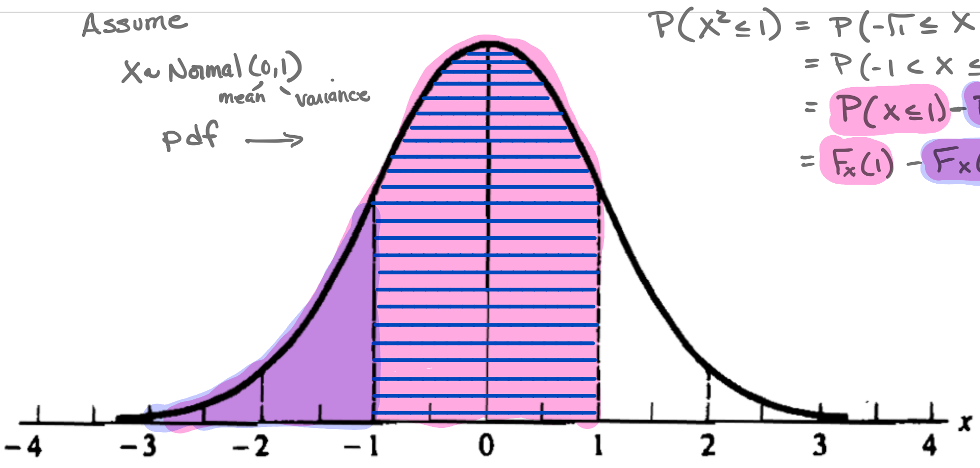
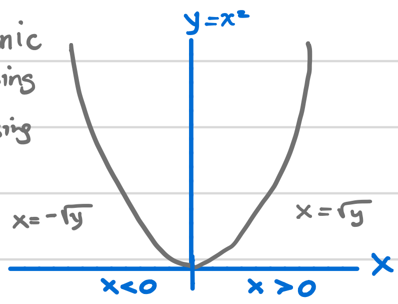


Figure 3.3.1. *Standard normal density*

If $Y = g(X) = X^2$
$F_Y(y) = P(Y \le y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$

not monotonic
decreasing $x<0$
increasing $x>0$



Find the pdf: $f_Y(y) = \frac{d}{dy} F_Y(y)$
$= \frac{d}{dy}\left[F_X(\sqrt{y}) - F_X(-\sqrt{y})\right]$
$= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})$ ← using chain rule

$f_Y(y) = \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right]$ ← pdf is sum of 2 pieces ⟩ intervals where $g(x)=x^2$ is monotone

Normal Distribution (continuous)

Normal(μ, σ^2)

pdf	$f(x\|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$, $\sigma > 0$ ↖ parameter space	sample space: $-\infty < x < \infty$, parameter space: $-\infty < \mu < \infty$,
mean and variance	$\mathrm{E}X = \mu$, $\mathrm{Var}\, X = \sigma^2$	
mgf	$M_X(t) = e^{\mu t + \sigma^2 t^2/2}$	
notes	Sometimes called the *Gaussian* distribution.	

Normal (0,1) $f(x \mid \mu=0, \sigma=1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2} \cdot I_{(-\infty,\infty)}(x)$

$I_{(-\infty,\infty)}(x) = \begin{cases} 1 & -\infty < x < \infty \\ 0 & \text{else} \end{cases}$

Theorem 2.1.8 *Let X have pdf $f_X(x)$, let $Y = g(X)$, and define the sample space $\mathcal{X}$ as in (2.1.7). Suppose there exists a partition, $A_0, A_1, \ldots, A_k$, of $\mathcal{X}$ such that $P(X \in A_0) = 0$ and $f_X(x)$ is continuous on each A_i. Further, suppose there exist functions $g_1(x), \ldots, g_k(x)$, defined on $A_1, \ldots, A_k$, respectively, satisfying*

i. $g(x) = g_i(x)$, *for* $x \in A_i$,

ii. $g_i(x)$ *is monotone on* A_i,

iii. *the set* $\mathcal{Y} = \{y\colon\ y = g_i(x)$ *for some* $x \in A_i\}$ *is the same for each* $i = 1, \ldots, k$, *and*

iv. $g_i^{-1}(y)$ *has a continuous derivative on* $\mathcal{Y}$, *for each* $i = 1, \ldots, k$.

Then

$$f_Y(y) = \begin{cases} \sum_{i=1}^k f_X\left(g_i^{-1}(y)\right)\left|\frac{d}{dy}g_i^{-1}(y)\right| & y \in \mathcal{Y} \\ 0 & \text{otherwise.} \end{cases}$$

Divide $\mathcal{X}$ into k Sets $A_1, \ldots, A_k$ s.t. $g(x)$ is monotone on each A_i.

- "We can ignore the A_0, since $P(X \in A_0) = 0$." ← handle interval endpoints.
- on each interval, $g_i(x)$ is one-to-one from $\mathcal{Y}$ onto A_i: $x = g^{-1}(y) \in A$ unique

Classic Example involving Normal (0,1) [N(0,1)] and chi-squared dist'n.

$$N(0,1) \quad f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} I_{(-\infty,\infty)}(x) \ ; \qquad I_{(-\infty,\infty)}(x) = \begin{cases} 1 & -\infty < x < \infty \\ 0 & \text{else} \end{cases}$$

Chi-Squared Distribution (continuous)

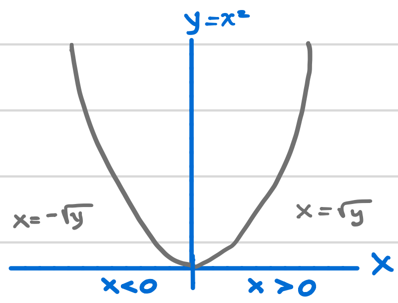
Chi squared(p)

pdf	$f(x\|p) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{(p/2)-1} e^{-x/2}; \quad 0 \le x < \infty; \quad p = 1, 2, \ldots$	sample space; parameters (degree of freedom)
mean and variance	$\mathrm{E}X = p, \quad \mathrm{Var}\,X = 2p$	$\Gamma(n) = (n-1)!$ if n is integer; $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$; $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ $\alpha > 0$
mgf	$M_X(t) = \left(\frac{1}{1-2t}\right)^{p/2}, \quad t < \frac{1}{2}$	
notes	Special case of the gamma distribution.	

$$X \sim \chi^2_p \qquad f(x|p) = \frac{1}{\Gamma(\frac{p}{2})2^{p/2}} x^{(p/2)-1} e^{-x/2} I_{(0,\infty)}(x), \qquad I_{(0,\infty)}(x) = \begin{cases} 1 & 0 \le x < \infty \\ 0 & \text{else} \end{cases}$$

greek letter
χ = Chi

Classic Example $X \sim N(0,1)$, $Y = X^2 \sim \chi^2_1$



$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} I_{(-\infty,\infty)}(x) \quad ; \quad y = x^2 = g(x)$$

$g(x) = x^2$ is monotone on $(-\infty,0)$ and $(0,\infty)$;

$x \to -\infty \quad y \to \infty \qquad x \to 0 \quad y \to 0$

$x \to 0 \quad y \to 0 \qquad x \to \infty \quad y \to \infty \Rightarrow \mathcal{Y} = (0,\infty)$

$A_0 = \{0\} \leftarrow$ interval endpoint $P(X \in A_0) = 0$

$A_1 = (-\infty,0), \quad g_1(x) = x^2, \quad g_1^{-1}(y) = -\sqrt{y}, \quad \left|\frac{d}{dy} g_i^{-1}(y)\right| = \frac{1}{2\sqrt{y}}$

$A_2 = (0,\infty), \quad g_2(x) = x^2, \quad g_2^{-1}(y) = \sqrt{y},$

$$\text{pdf} \quad f_Y(y) = \left(\frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{y})^2/2}\left(\frac{1}{2\sqrt{y}}\right) + \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2}\left(\frac{1}{2\sqrt{y}}\right)\right) * I_{(0,\infty)}(y)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-y/2}\left(\frac{1}{2\sqrt{y}} + \frac{1}{2\sqrt{y}}\right) I_{(0,\infty)}(y)$$

$$= \frac{1}{\sqrt{2\pi}} \frac{e^{-y/2}}{\sqrt{y}} I_{(0,\infty)}(y)$$

Which we will soon recognize as a χ^2_1

Chi-Squared Distribution with 1 degree of freedom

$$f(y \mid p=1) = \frac{1}{\Gamma(1/2)\sqrt{2}} y^{(1/2-1)} e^{-y/2} I_{(0,\infty)}(y)$$

$$= \frac{1}{\sqrt{2\pi}} \frac{e^{-y/2}}{\sqrt{y}} I_{(0,\infty)}(y)$$

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

* see Appendix

//

Finally

One of most useful transformations : (popular Q-exam topic)

Theorem 2.1.10 (Probability integral transformation) *Let X have continuous cdf $F_X(x)$ and define the random variable Y as $Y = F_X(X)$. Then Y is uniformly distributed on $(0,1)$, that is, $P(Y \le y) = y, 0 < y < 1$.*

One of most useful transformations :

Theorem 2.1.10 (Probability integral transformation) *Let X have continuous cdf $F_X(x)$ and define the random variable Y as $Y = F_X(X)$. Then Y is uniformly distributed on $(0,1)$, that is, $P(Y \le y) = y, 0 < y < 1$.*

X continuous w/ cdf $F_X(x)$

$Y = F_X(x)$

$Y \sim U(0,1)$ $\quad P(Y \le y) = y, \ 0 < y < 1.$

Example: Assume $X \sim$ exponential

$$f_X(x) = \frac{1}{\lambda} e^{-x/\lambda} I_{(0,\infty)}^{(x)}$$

$$F_X(x) = \int_0^x \frac{1}{\lambda} e^{-t/\lambda} dt$$

$$= 1 - e^{-x/\lambda}$$

[let $u = -t/\lambda$

$\frac{du}{dt} = -1/\lambda \quad du = -\frac{1}{\lambda} dt$

$\int_0^x \frac{1}{\lambda} e^{-t/\lambda} dt = \int -e^u du = -e^u \Big|$

$= -e^{-t/\lambda} \Big|_0^x = -e^{-x/\lambda} + 1$]

Let $Y = 1 - e^{-x/\lambda}$

Find dist'n of $Y = F_X(x)$

- Show monotonic ft'n:

$$\frac{d}{dx}(1 - e^{-x/\lambda}) = \frac{e^{-x/\lambda}}{\lambda} > 0 \Rightarrow \text{monotonic}$$

$\left[\frac{d}{dx}(-e^{-x/\lambda}) = (-)(-\frac{1}{\lambda}) e^{-x/\lambda}\right]$

- Find $g^{-1}(y)$: Solve for x in terms of y

$$x = -\lambda \log(1-y) = g^{-1}(y)$$

[$e^{-x/\lambda} = 1 - y$

$-\frac{x}{\lambda} = \log(1-y)$

$x = -\lambda \log(1-y)$]

- Determine sample space, $\mathcal{Y}$

$$0 \le Y \le 1$$

[If $x = 0$ $\quad y = 0$

$x \to \infty$ $\quad y \to 1$]

- Calculate $\frac{d}{dy}(g^{-1}(y))$ (derivate old wrt new)

$$\frac{d(g^{-1}(y))}{dy} = \frac{d}{dy}(-\lambda \log(1-y))$$

[$\frac{d}{dy}(-\lambda \log(1-y))$

$= \frac{-\lambda}{(1-y)}(-1) = \frac{\lambda}{1-y}$]

- In $f_X(x)$ replace x by $g^{-1}(y)$ multiply by $\left|\frac{d}{dy} g^{-1}(y)\right|$ and identify $\mathcal{Y}$

pdf of Y $\Rightarrow f_Y(y) = 1 \, I_{(0,1)}^{(y)}$

$$Y \sim U(0,1)$$

[$f_Y(y) = \frac{1}{\lambda} \exp\{-\cancel{\lambda} \log(1-y)/\cancel{\lambda}\} \Big| * \Big| \frac{\lambda}{1-y} \Big|$

$= \frac{1}{\lambda}(1-y) \Big| \frac{\lambda}{1-y} \Big| = 1$]

Go Backwards: Find dist'n $X = \lambda \log(1-Y) = F_X^{-1}(Y)$

If $Y \sim U(0,1)$, Find dist'n of $X = -\lambda \log(1-Y)$

switch $\overset{\text{new}}{X} = -\lambda \log(1-\overset{\text{old}}{Y})$

- Show monotonic ft'n:

$$\frac{d}{dy}(-\lambda \log(1-y)) = \frac{\lambda}{1-y} > 0 \quad \text{for } 0<y<1$$

[previous page $\frac{d}{dy}(-\lambda \log(1-y)) = \frac{d}{dy} g^{-1}(y)$

- Find $g^{-1}(x)$: Solve for y in terms of x

$$y = 1 - e^{-x/\lambda}$$

[$y = F_X(x)$ previous page

- Determine sample space, $\mathcal{X}$

$$0 \le X < \infty$$

[As expected if $y=0$ $x=0$
$y \to 1$ $x \to \infty$

- Calculate $\frac{d}{dx}(g^{-1}(x))$

$$\frac{d}{dx}(1-e^{-x/\lambda}) = \frac{e^{-x/\lambda}}{\lambda}$$

[previous page $\frac{d}{dx} g(x)$

- In $f_Y(y)$ replace y by $g^{-1}(x)$ multiply by $\left|\frac{d}{dx} g^{-1}(x)\right|$ and identify $\mathcal{X}$

$$f_X(x) = 1 * \left|\frac{e^{-x/\lambda}}{\lambda}\right| * I_{(0,\infty)}(x)$$

$X \sim$ exponential (λ)

Why so important? If you can generate a $U(0,1)$, you can generate random variables from any continuous dist'n.

To generate exp (λ) $(n=10{,}000)$

generate 10,000 values of $Y \sim U(0,1)$

calculate $X = \lambda \log(1-Y)$

$\Rightarrow$ 10,000 exp (λ).

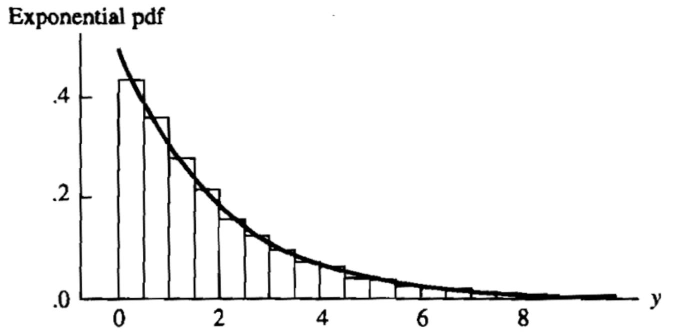


Figure 5.6.1. *Histogram of* 10,000 *observations from an exponential pdf with* $\lambda = 2$, *together with the pdf*

§2.2 Expected Values (average value from dist'n)

Definition 2.2.1 The *expected value* or *mean* of a random variable $g(X)$, denoted by $\mathrm{E}\, g(X)$, is

$$\mathrm{E}\, g(X) = \begin{cases} \int_{-\infty}^{\infty} g(x) f_X(x)\, dx & \text{if } X \text{ is continuous} \\ \sum_{x \in \mathcal{X}} g(x) f_X(x) = \sum_{x \in \mathcal{X}} g(x) P(X = x) & \text{if } X \text{ is discrete,} \end{cases}$$

provided that the integral or sum exists. If $\mathrm{E}|g(X)| = \infty$, we say that $\mathrm{E}\, g(X)$ does not exist. (Ross 1988 refers to this as $\underbrace{\text{the "law of the unconscious statistician."}}_{\text{LOTUS - see appendix.}}$ We do not find this amusing.)

$$E[X] = \begin{cases} \int_{-\infty}^{\infty} x f_X(x)\, dx & X \text{ is continuous} \\ \sum_{x \in \mathcal{X}} x f_X(x)\, dx & X \text{ is discrete} \end{cases}$$

weighted average over values of X.

- Note $E[X]$ will be a function of dist'n parameters not a function of x.

Example 2.2.2 (Exponential mean)

$X \sim \text{exponential}(\lambda)$ $\qquad f_X(x) = \frac{1}{\lambda} e^{-x/\lambda} I_{[0,\infty)}(x), \quad \lambda \geq 0$

$$E[X] = \int_0^\infty x \frac{e^{-x/\lambda}}{\lambda} dx$$
$$= -x e^{-x/\lambda} \Big|_0^\infty + \int_0^\infty e^{-x/\lambda} dx$$
$$= 0 + \int_0^\infty e^{-x/\lambda}$$
$$= -\lambda e^{-x/\lambda} \Big|_0^\infty = 0 - [-\lambda] = \lambda$$

Integration by parts:
$\int_a^b u\,dv = uv\big|_a^b - \int_a^b v\,du$

$u = x \qquad dv = e^{-x/\lambda}/\lambda$

$du = dx \qquad v = -e^{-x/\lambda}$

change of variables $\quad u = -x/\lambda$
$\quad dx = -\lambda du$

Example 2.2.3 (Binomial mean)

$X \sim \text{bin}(n,p) \qquad P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} * I_{[0,1,\dots,n]}(x)$

$$E[X] = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$
$$= \sum_{x=1}^{n} n \underbrace{\binom{n-1}{x-1}}_{\text{appendix}} p^x (1-p)^{n-x} = \sum_{y=0}^{n-1} n \binom{n-1}{y} p^{y+1} (1-p)^{n-(y+1)}$$
$$= np \underbrace{\sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y}}_{\text{bin}(n-1,p) \text{ must sum to } 1} = np$$

(details later...)

change of variables
$y = x-1 \rightarrow x = y+1$
$x=1 \rightarrow y=0;$
$x=n \rightarrow y=n-1$

Appendix

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

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty t^{1/2-1} e^{-t}\,dt = \sqrt{\pi}$$

$$= \int_0^\infty \frac{e^{-t}}{\sqrt{t}}\,dt$$

let $u=\sqrt{t}$
$t = u^2$

$\frac{du}{dt} = \frac{1}{2} t^{-1/2}$

$du = \frac{1}{2} t^{-1/2}\,dt$

$$= \int_0^\infty 2e^{-u^2}\,du$$

$$= \sqrt{\pi}\int_0^\infty \frac{2e^{-u^2}}{\sqrt{\pi}}\,du$$

$$= \sqrt{\pi}\left(2\int_0^\infty \frac{e^{-u^2}}{\sqrt{\pi}}\,du\right)$$

$$= \sqrt{\pi} \quad //$$

$X \sim N(0, \frac{1}{2})$ $\quad f_X(x) = \frac{1}{\sqrt{2\pi}\left(\frac{1}{\sqrt{2}}\right)} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\frac{1}{2}}\right)\right)$

$$= \underbrace{\frac{1}{\sqrt{\pi}} e^{-x^2} I_{(-\infty,\infty)}(x)}_{\text{symmetric about } 0.}$$

pdfs integrate to 1.
(we will show this later.)

= Show $x\binom{n}{x} = n\binom{n-1}{x-1}$

$$\cancel{x}\left(\frac{n(n-1)!}{\cancel{x}(x-1)!} * \frac{1}{(n-x)!}\right), \quad \text{since } (n-1)-(x-1) = n-x$$

$$= n\binom{n-1}{x-1}$$

MS Theory
Appendix

Lecture-5
13

* Note on def'n 2.2.1 (Expectation)
Law of the unconscious statistician (LOTUS)

iPad 2:35 PM 62%

Law of the unconscious statistician - Wikipedia
en.m.wikipedia.org

WIKIPEDIA Search Wikipedia

Law of the unconscious statistician (LOTUS)

In probability theory and statistics, the **law of the unconscious statistician** (sometimes abbreviated LOTUS) is a theorem used to calculate the expected value of a function $g(X)$ of a random variable X when one knows the probability distribution of X but one does not explicitly know the distribution of $g(X)$.

The form of the law can depend on the form in which one states the probability distribution of the random variable X. If it is a discrete distribution and one knows its probability mass function f_X (but not $f_{g(X)}$), then the expected value of $g(X)$ is

$$\mathrm{E}[g(X)] = \sum_x g(x) f_X(x),$$

where the sum is over all possible values x of X. If it is a continuous distribution and one knows its probability density function f_X (but not $f_{g(X)}$), then the expected value of $g(X)$ is

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x$$

(provided the values of X are real numbers as opposed to vectors, complex numbers, etc.).

Regardless of continuity-versus-discreteness and related issues, if one knows the cumulative probability distribution function F_X (but not $F_{g(X)}$), then the expected value of $g(X)$ is given by a Riemann–Stieltjes integral

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x)\,\mathrm{d}F_X(x)$$

(again assuming X is real-valued).[1][2]

The above equation is sometimes known as the law of the unconscious statistician, as statisticians have been accused of using the identity without realizing that it must be treated as the result of a rigorously proved theorem, not merely a definition.[3]

However, the result is so well known that it is usually used without stating a name for it: the name is not extensively used. For justifications of the result for discrete and continuous random variables see.[4]

Definition 2.2.1 The *expected value* or *mean* of a random variable $g(X)$, denoted by $\mathrm{E}\,g(X)$, is

$$\mathrm{E}\,g(X) = \begin{cases} \int_{-\infty}^{\infty} g(x) f_X(x)\,dx & \text{if } X \text{ is continuous} \\ \sum_{x\in\mathcal{X}} g(x) f_X(x) = \sum_{x\in\mathcal{X}} g(x) P(X=x) & \text{if } X \text{ is discrete,} \end{cases}$$

provided that the integral or sum exists. If $\mathrm{E}|g(X)| = \infty$, we say that $\mathrm{E}\,g(X)$ does not exist. (Ross 1988 refers to this as the "law of the unconscious statistician." We do not find this amusing.)