MS Theory - I

Lecture 6

Review Continuous Dist'ns

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 $\mathcal{Y} = \{y : y = g(x)$ for some $x \in \mathcal{X}\}$

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*

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

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

Approach:

Thm 2.1.5

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

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: $f_Y(y)$

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

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

Review

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 $u = -t/\lambda$

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

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

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

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

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

- Show monotonic ft'n:

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

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

- 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 $\Rightarrow f_Y(y) = 1 \, I_{(0,1)}(y)$

$Y \sim U(0,1)$

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

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

Review § 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 the "law of the unconscious statistician." We do not find this amusing.)

LOTUS - see appendix.

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

weighted average over values of X

- Note this will be a function of dist'n parameters not a function of x.

Example 2.2.2 (Exponential mean)

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

$$E[X] = \int_0^{\infty} x \frac{e^{-x/\lambda}}{\lambda} dx$$

$$= -xe^{-x/\lambda}\Big|_0^{\infty} + \int_0^{\infty} e^{-x/\lambda} dx$$

$$= 0 + \int_0^{\infty} e^{-x/\lambda}$$

$$= -\lambda e^{-x/\lambda}\Big|_0^{\infty} = 0 - [-\lambda] = \lambda$$

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

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

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

Example 2.2.3 (Binomial mean)

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

$$E[X] = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} \underbrace{n \binom{n-1}{x-1}}_{\text{appendix}} p^x (1-p)^{n-x} = \sum_{y=0}^{n-1} n \binom{n-1}{y} p^{y+1} (1-p)^{n-(y+1)}$$

$$= np \underbrace{\sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y}}_{\text{bin}(n-1,p) \text{ must sum to } 1} = np$$

(details later...)

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

Example 2.2.4 (Cauchy mean) → Classic Example
E[X] does not exist.

Cauchy(θ,σ)	
pdf	$f(x\|\theta,\sigma)=\frac{1}{\pi\sigma}\frac{1}{1+\left(\frac{x-\theta}{\sigma}\right)^2},\quad -\infty<x<\infty;\quad -\infty<\theta<\infty,\quad \sigma>0$
mean and variance	do not exist
mgf	does not exist
notes	Special case of Student's t, when degrees of freedom $=1$. Also, if X and Y are independent $\mathrm{n}(0,1)$, X/Y is Cauchy.

Cauchy $(\sigma=1)$

$$f_X(x)=\frac{1}{\pi}\frac{1}{1+x^2}*I_{(-\infty,\infty)}(x)$$

$$E[|X|]=\int_{-\infty}^{\infty}\frac{|x|}{\pi}\frac{1}{1+x^2}dx=\frac{2}{\pi}\int_0^{\infty}\frac{x}{1+x^2}dx=\frac{2}{\pi}\frac{\log(1+x^2)}{2}\Big|_0^{\infty}$$

$$=\lim_{M\to\infty}\frac{\log(1+M^2)}{\pi}=\infty$$

E[X] does not exist for the Cauchy dist'n.

Properties of Expectation

Theorem 2.2.5 *Let X be a random variable and let a, b, and c be constants. Then for any functions $g_1(x)$ and $g_2(x)$ whose expectations exist,*

Proof in C&B pg. 57

a. $\mathrm{E}(ag_1(X)+bg_2(X)+c)=a\mathrm{E}\,g_1(X)+b\mathrm{E}\,g_2(X)+c$.
b. *If $g_1(x)\ge 0$ for all x, then $\mathrm{E}\,g_1(X)\ge 0$.*
c. *If $g_1(x)\ge g_2(x)$ for all x, then $\mathrm{E}\,g_1(X)\ge \mathrm{E}\,g_2(X)$.*
d. *If $a\le g_1(x)\le b$ for all x, then $a\le \mathrm{E}\,g_1(X)\le b$.*

$E[aX+b]=aE[X]+b$; a & b constants

$E[X-E[X]]=0$

example $X\sim \mathrm{bin}(n,p)$

$E[X]=np$

$E[X-np]=E[X]-np=0$

If we are interested in $E[g(X)]$, where g is some nonlinear ft'n.

2 ways: 1) $E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$.

2) Transformation $Y = g(X)$

Find dist'n of $Y = g(X)$, $f_Y(y)$.

$E[g(X)] = E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy$.

Example (Uniform - exponential).

LOTUS * lecture 5

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

- Assume $f_X(x) = 1 * I_{[0,1]}(x)$ $\quad X \sim U(0,1)$

$g(X) = -\log(X)$

$E[g(X)] = E[-\log(X)]$

$= \int_0^1 -x \log(x)\, dx$

$= -\left[x \log(x)\Big|_0^1 - \int_0^1 \frac{x}{x} dx\right]$

$= 0 + x\Big|_0^1 = 1$

integration by parts.
$\int u\,dv = vu| - \int v\,du$

$u = \log(x) \quad dv = 1$
$du = \frac{1}{x} dx \quad v = x$

$Y = g(X)$

$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy$

- Assume $Y = g(X) = -\log(X)$

$f_Y(y) = e^{-y} \sim$ exponential (1)

$E[Y] = 1$

Example: details Lecture 4 and 5 (Review)

example page 3.

§2.3 Moments and Moment Generating Functions

Definition 2.3.1 For each integer n, the nth *moment of* X (or $F_X(x)$), μ'_n, is

$$\mu'_n = \mathrm{E}\, X^n.$$

The nth *central moment of* X, μ_n, is

$$\mu_n = \mathrm{E}(X - \mu)^n,$$

where $\mu = \mu'_1 = \mathrm{E}\, X$.

	Moments	Central Moment
$X \sim F_X(x)$	$\mu'_1 = E[X] = \mu$	$\mu_1 = E[X-\mu] = E[X-E[X]] = 0$
	$\mu'_2 = E[X^2]$	$\mu_2 = E[(X-\mu)^2] = Var(X)$
	$\vdots$	$\vdots$
	$\mu'_n = E[X^n]$	$\mu_n = E[(X-\mu)^n$

$E[x]$ = mean
$Var[x] = E[(x-\mu)^2]$

Definition 2.3.2 The *variance* of a random variable X is its second central moment, $\text{Var}\, X = \text{E}(X - \text{E}\, X)^2$. The positive square root of $\text{Var}\, X$ is the *standard deviation* of X.

- Variance measure of spread
- Large Variance $\Rightarrow$ X more variable
- $Var(x) = 0$ $\quad E[(x - E[x])^2] = 0$
  $x = E[x]$ with prob. 1 (no variation in X)
- Small $Var(x) \Rightarrow$ X less variable and close to $E[x]$ = mean
- If X has units: age, height, ... $\quad Var(x)$: Units^2
- Standard Deviation $= \sqrt{Var(x)}$ (same units as X)

Example 2.3.3 (Exponential variance)

$X \sim$ exponential (λ)

$E[x] = \lambda$ Example: page 3

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

$$Var(x) = E[(x-\lambda)^2] = \int_0^\infty (x-\lambda)^2 \frac{1}{\lambda} e^{-x/\lambda} dx$$

$$= \int_0^\infty (x^2 - 2x\lambda + \lambda^2) \frac{1}{\lambda} e^{-x/\lambda} dx$$

$$= \underbrace{\int_0^\infty \frac{x^2}{\lambda} e^{-x/\lambda} dx}_{\substack{2\lambda^2 \\ \text{see appendix (i)}}} \underbrace{-2\lambda \int_0^\infty \frac{x}{\lambda} e^{-x/\lambda} dx}_{-2\lambda E[x] = -2\lambda^2} + \underbrace{\int_0^\infty \lambda e^{-x/\lambda} dx}_{\substack{\lambda^2 \\ \text{see appendix (ii)}}}$$

$$2\lambda^2 - 2\lambda^2 + \lambda^2 = \lambda^2 \quad //$$

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

$$\left.\begin{aligned} E[X] &= \lambda \\ Var[X] &= \lambda^2 \end{aligned}\right\} \text{ Function of parameter}$$

Var(x) smaller, smaller values λ.

$f_X(x)$

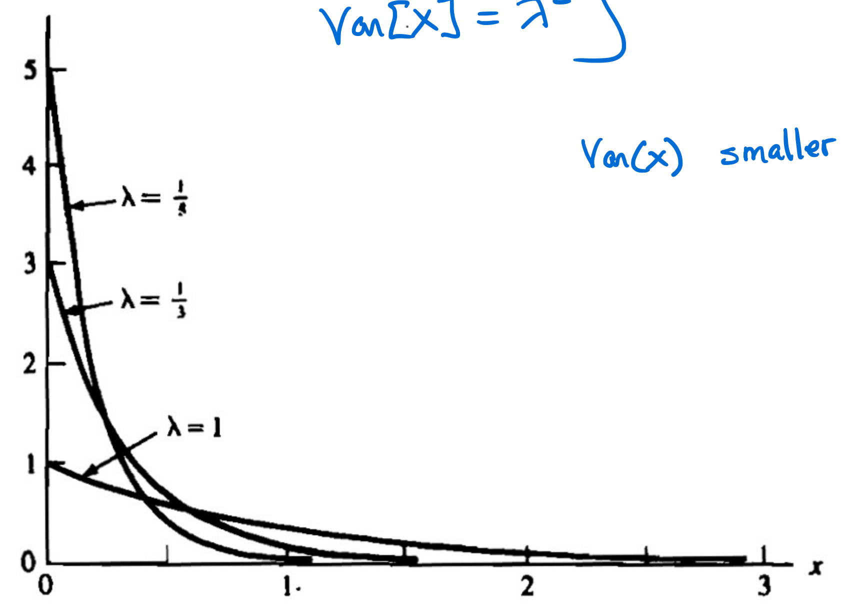

Figure 2.3.1. *Exponential densities for* $\lambda = 1, \frac{1}{3}, \frac{1}{5}$

Theorem 2.3.4 *If X is a random variable with finite variance, then for any constants a and b,*

$$\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}\,X.$$

Proof: $\mathrm{Var}(aX+b) = E\left[\left((aX+b) - E[(aX+b)]\right)^2\right]$

$$= E\left[\left((aX+\cancel{b}) - (aE[X] + \cancel{b})\right)^2\right]$$

$$= E\left[\left(a(X) - aE[X]\right)^2\right]$$

$$= a^2 E\left[(X - E[X])^2\right]$$

$$= a^2\,\mathrm{Var}(X) \quad //$$

Easier Calculation

$$\mathrm{Var}(X) = E\left[(X - E[X])^2\right] = E[X^2] - (E[X])^2$$

Since: $\mathrm{Var}(X) = E\left[(X - E(X))^2\right]$

$$= E\left[X^2 - 2XE[X] + (E[X])^2\right]$$

Note $E[X]$ is a constant

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

If b is constant $E[b] = b$

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

Example 2.3.5 (Binomial Variance)

$X \sim bin(n, p)$ $\quad f_X(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} \cdot I_{[0,1,2,\dots n]}(x)$

Know $E[X] = np$ (page 3)

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

Calculate $E[X^2] = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x}$

aside

$$x^2 \binom{n}{x} = x^{\cancel{2}} \frac{n(n-1)!}{\cancel{x}(x-1)!(n-x)!} = xn \binom{n-1}{x-1}$$

$$E[X^2] = n \sum_{x=1}^{n} x \binom{n-1}{x-1} p^x (1-p)^{n-x}$$

change of variables

$y = x-1$
$x = y+1$

$$= n \sum_{y=0}^{n-1} (y+1) \binom{n-1}{y} p^{y+1} (1-p)^{n-1-y}$$

$$= np \underbrace{\sum_{y=0}^{n-1} y \binom{n-1}{y} p^y (1-p)^{n-1-y}}_{E[y],\ y \sim bin(n-1, p)}$$

$$+ np \underbrace{\sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y}}_{=1 \text{ since pdf sum } bin(n-1,p)}$$

$$= np((n-1)p) + np$$

$$= n(n-1)p^2 + np$$

$$Var(x) = n(n-1)p^2 + np - (np)^2$$

$$= \cancel{n^2p^2} - np^2 + np - \cancel{n^2p^2}$$

$$= np(1-p).$$

Can continue to higher order moments
- sometime interested in 3rd or 4th order moments
- rarely higher order moments of interest.

Future : Where to next ?

Moment Generating Functions

Definition 2.3.6 Let X be a random variable with cdf F_X. The *moment generating function (mgf) of* X (or F_X), denoted by $M_X(t)$, is

$$M_X(t) = \mathrm{E}\, e^{tX},$$

provided that the expectation exists for t in some neighborhood of 0. That is, there is an $h > 0$ such that, for all t in $-h < t < h$, $\mathrm{E}e^{tX}$ exists. If the expectation does not exist in a neighborhood of 0, we say that the moment generating function does not exist.

More explicitly, we can write the mgf of X as

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx \qquad \text{if } X \text{ is continuous,}$$

or

$$M_X(t) = \sum_x e^{tx} P(X = x) \qquad \text{if } X \text{ is discrete.}$$

It is very easy to see how the mgf generates moments. We summarize the result in the following theorem.

Theorem 2.3.7 *If X has mgf $M_X(t)$, then*

$$\mathrm{E}\, X^n = M_X^{(n)}(0),$$

where we define

$$M_X^{(n)}(0) = \frac{d^n}{dt^n} M_X(t)\Big|_{t=0}.$$

That is, the nth moment is equal to the nth derivative of $M_X(t)$ evaluated at $t = 0$.

Appendix

i) Calculate $\int \frac{x^2 e^{-x/\lambda}}{\lambda} dx$

$= \frac{1}{\lambda}\left[\int x^2 e^{-x/\lambda} dx\right]$

$= \frac{1}{\lambda}\left[\int (u\lambda)^2 e^{-u} \lambda du\right] = \frac{1}{\lambda}\left[\lambda^3 \int u^2 e^{-u} du\right]$

Let $u = x/\lambda$ $\quad x = u\lambda$

$\frac{du}{dx} = \frac{1}{\lambda}$ $\quad dx = \lambda du$

$dx = \lambda du$ $\quad$ $\left[\begin{array}{l} x=0 \rightarrow u=0 \\ x\rightarrow\infty \rightarrow u\rightarrow\infty \end{array}\right.$

Solve: $\int u^2 e^{-u} du$

$= \underbrace{-u^2 e^{-u}\Big|_0^\infty}_{0} - \int_0^\infty -2u e^{-u} du$

$= 2 \int u e^{-u} du$

use integration by parts

$\int fg' = fg - \int f'g$

$f = u^2 \quad g' = e^{-u}$

$f' = 2u \quad g = -e^{-u}$

$\int fg' = -u^2 e^{-u}\Big|_0^\infty - \int_0^\infty 2u(-e^{-u}) du$

Solve $\int u e^{-u} du$

$= \underbrace{u e^{-u}\Big|_0^\infty}_{0} + \int_0^\infty e^{-u} du$

use integration by parts

$\int fg' = fg - \int f'g$

$f = u \quad g' = e^{-u}$

$f' = 1 \quad g = -e^{-u}$

$\int fg' = -u e^{-u}\Big|_0^\infty - \int -e^{-u} du$

Solve $\int_0^\infty e^{-u} du$

$= -e^{-u}$

$$\int_0^\infty \frac{x^2 e^{-x/\lambda}}{\lambda} dx = \lambda^2\left[-2\left[e^{-x/\lambda}\right]_0^\infty\right]$$

$$= \lambda^2[-2(0-1)] = 2\lambda^2$$

---

ii) Calculate $\int_0^\infty \lambda e^{-x/\lambda} dx$

$= \lambda \int e^{-x/\lambda} dx$

$= \lambda^2$

Transformation:

$u = -x/\lambda \quad dx = -\lambda du$

(above)

$\int_0^\infty e^{-x/\lambda} dx = \int e^{+u} \lambda du$

$= \lambda e^{u}\Big|_0^\infty = \lambda$

Note: $\int_0^\infty e^{-x/\lambda} dx = \lambda \int_0^\infty \frac{e^{-x/\lambda}}{\lambda} dx$

$= \lambda * 1$

since pdf must $\int$ to 1.