Review

Definition 4.1.1 An *n-dimensional random vector* is a function from a sample space S into $\Re^n$, n-dimensional Euclidean space.

bivariate: $(X(s), Y(s)) \in \mathbb{R}^2$ or $(X,Y) \in \mathbb{R}^2$

multivariate $(X_1(s), X_2(s), \ldots X_n(s)) \in \mathbb{R}^n$
or $(X_1, X_2, \ldots X_n) \in \mathbb{R}^n$

"Classic" Example:
Roll 2 fair 6-sided dice (36 possible outcomes, all equally likely)

$$S = \begin{bmatrix} 1,1 & 1,2 & 1,3 & \boxed{1,4} & 1,5 & 1,6 \\ 2,1 & 2,2 & 2,3 & 2,4 & \boxed{2,5} & 2,6 \\ 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \end{bmatrix}$$

X = sum of dice $\mathcal{X} = (2, 3, \ldots 12)$
Y = | difference of 2 dice |. $\mathcal{Y} = (0, 1, \ldots 5)$

X=7, Y=0

	x: 2	3	4	5	6	7	8	9	10	11	12	$f_Y(y)$
y = 0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$	$\boxed{0}$	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$	$\frac{3}{18}$
1		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$5/18$
2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$			$4/18$
3				$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$				$3/18$
4					$\frac{1}{18}$		$\frac{1}{18}$					$2/18$
5						$\frac{1}{18}$						$1/18$
$f_X(x)$:	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$	1

Table 4.1.1. *Values of the joint pmf* $f(x,y)$

What is $Pr(X=6 + Y=0)$? $= P(\text{Roll } (3,3)) = 1/36$
$Pr(X=5 + Y=3)$? $= P((4,1) \text{ or } (1,4)) = 2/36 = 1/18$
$P(X=7 + Y=3)$? $= P((5,2) \text{ or } (2,5)) = 2/36 = 1/18$
$P(X=7 + Y=0)$? $= P(\emptyset) = 0$

$P(X=7 + Y \leq 4)$? $= P((4,3) \text{ or } (3,4) \text{ or } (5,2) \text{ or } (2,5)) = \frac{4}{36} = \frac{1}{9}$

Review
Joint pmf:

Definition 4.1.3 Let (X, Y) be a discrete bivariate random vector. Then the function $f(x, y)$ from $\Re^2$ into $\Re$ defined by $f(x,y) = P(X = x, Y = y)$ is called the *joint probability mass function* or *joint pmf* of (X, Y). If it is necessary to stress the fact that f is the joint pmf of the vector (X, Y) rather than some other vector, the notation $f_{X,Y}(x, y)$ will be used.

$$P((x,y)\in A) = \sum_{(x,y)\in A} f(x,y)\ , \quad A \text{ any subset of } \mathbb{R}^2$$

$$E[g(X,Y)] = \sum_{(x,y)\in \mathbb{R}^2} g(x,y) f(x,y)$$

$$\sum_{(x,y)\in \mathbb{R}^2} f(x,y) = P((X,Y)\in \mathbb{R}^2) = 1$$

Theorem 4.1.6 *Let (X, Y) be a discrete bivariate random vector with joint pmf $f_{X,Y}(x, y)$. Then the marginal pmfs of X and Y, $f_X(x) = P(X = x)$ and $f_Y(y) = P(Y = y)$, are given by*

$$f_X(x) = \sum_{y\in\Re} f_{X,Y}(x,y) \quad \text{and} \quad f_Y(y) = \sum_{x\in\Re} f_{X,Y}(x,y).$$

$$\text{Sample Space} \underset{\cancel{\leftarrow}}{\longrightarrow} f_{xy}(x,y) \underset{\cancel{\leftarrow}}{\longrightarrow} f_x(x),\ f_y(y)$$

Review

Continuous Random Vectors

Definition 4.1.10 A function $f(x, y)$ from $\Re^2$ into $\Re$ is called a *joint probability density function* or *joint pdf of the continuous bivariate random vector* (X, Y) if, for every $A \subset \Re^2$,

$$P((X,Y) \in A) = \underbrace{\int_A \int}_{} f(x,y)\, dx\, dy.$$

double integrals, integrate over all $(x,y) \in A$.

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

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

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

Function satisfying $f(x,y) \geq 0 \ \forall \ (x,y) \in \mathbb{R}^2$

$\&$ $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dx\, dy = 1.$

is joint pdf of continuous random vector (X,Y).

Example: Calculating joint probabilities - I

$$f(x,y) = 6xy^2 I_{(0,1)}^{(x)} I_{(0,1)}^{(y)}$$

- Check pdf: Clearly $f(x,y) \geq 0$ for $0<x<1$ & $0<y<1$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dx\, dy = \int_0^1\int_0^1 6xy^2\, dx\, dy = \int_0^1 3x^2y^2\Big|_0^1 dy$$

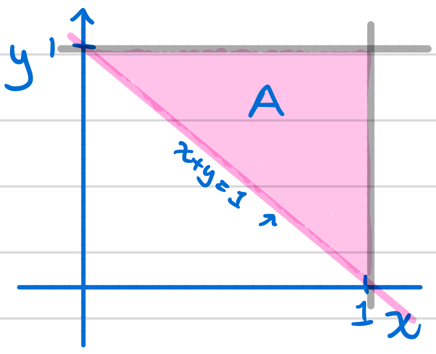
$$= \int_0^1 3y^2\, dy = y^3\Big|_0^1 = 1$$

Review

Example cont. $f(x,y) = 6xy^2 I_{(0,1)}^{(x)} I_{(0,1)}^{(y)}$

Calculate: $P(X+Y \geq 1)$ let $A = \{(x,y): x+y \geq 1\}$

$= P((X,Y) \in A)$



$A = \{(x,y): x+y \geq 1, 0<x<1, 0<y<1\}$

$= \{(x,y): x \geq 1-y, 0<x<1, 0<y<1\}$

$= \{(x,y): 1-y \leq x < 1, 0<y<1\}$

$$P(X+Y \geq 1) = \int_A\int f(x,y)\,dx\,dy = \int_0^1 \int_{1-y}^1 6xy^2\,dx\,dy = \int_0^1 3x^2y^2\Big|_{1-y}^1 dy$$

$$= \int_0^1 3y^2[1^2 - (1-y)^2]\,dy = \int_0^1 3y^2[\cancel{1} - (\cancel{1} - 2y + y^2)]dy = \int_0^1 6y^3 - 3y^4\,dy$$

$$= \frac{6y^4}{4} - \frac{3y^5}{5}\Big|_0^1 = \frac{3}{2}y^4 - \frac{3}{5}y^5\Big|_0^1 = -\frac{3y^4(2y-5)}{10}\Big|_0^1 = \frac{-3(2-5)}{10} = \frac{9}{10}$$

Calculate $f_X(x)$ $\quad f_X(x) = \int_y f(x,y)\,dy = \int_0^1 6xy^2\,dy$

$$= 2xy^3\Big|_0^1 = 2x \quad 0<x<1$$

$$f_X(x) = 2x\, I_{(0,1)}^{(x)}$$

Check $f_X(x)$ is a pdf

$f_X(x) \geq 0$ for $0<x<1$

$\int_0^1 2x\,dx = x^2\Big|_0^1 = 1$

Calculate $P(\frac{1}{2} < X < \frac{3}{4}) = \int_{1/2}^{3/4} 2x\,dx = x^2\Big|_{1/2}^{3/4}$

$$= \frac{9}{16} - \frac{1}{4} = \frac{9}{16} - \frac{4}{16} = \frac{5}{16}$$

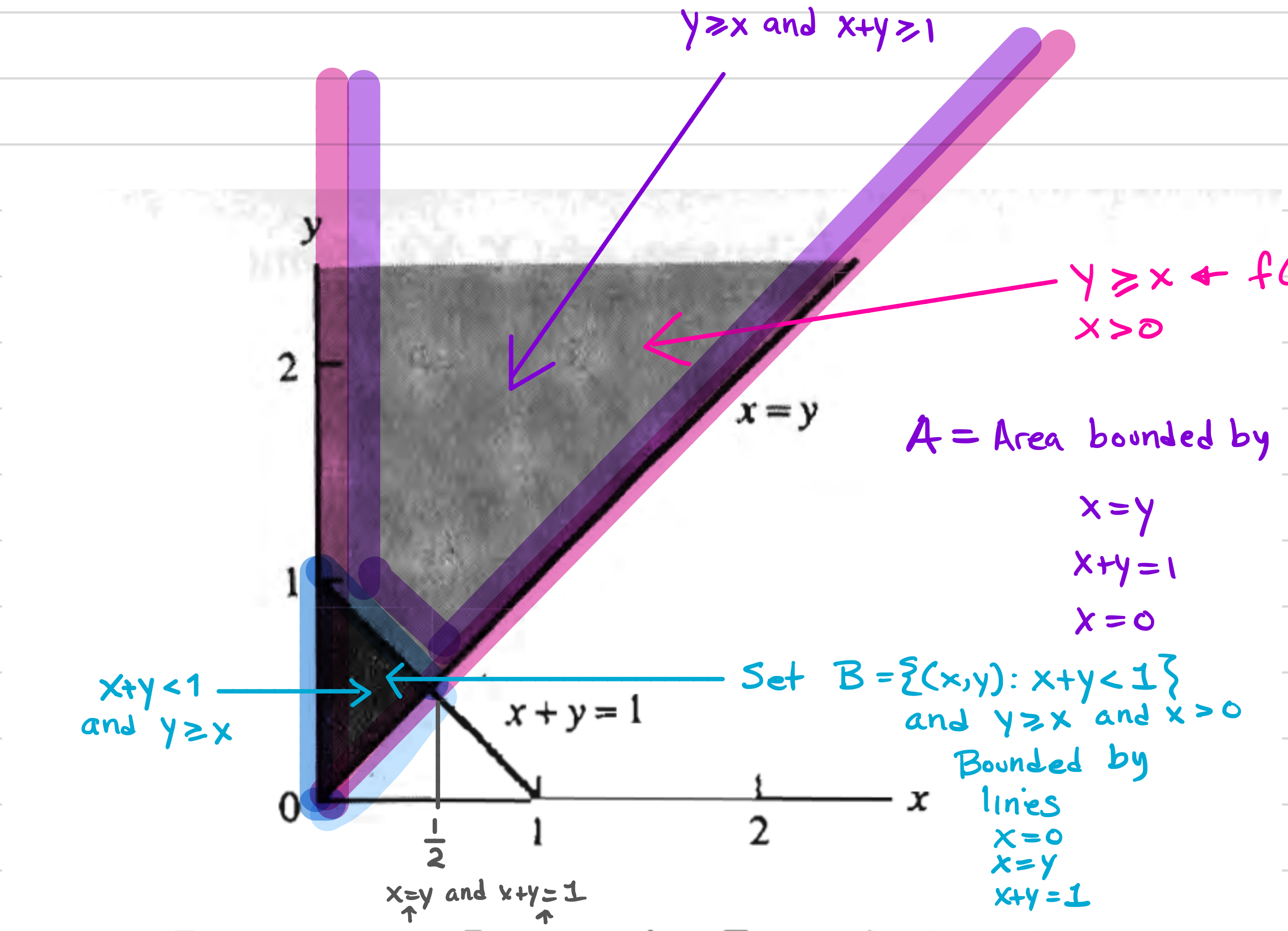
Example: Calculating joint probabilities

$$f(x,y) = \underset{\text{not fct'n of } x}{\underline{e^{-y}}}\ \underset{\text{fct'n } x}{\underline{I^{(x,y)}_{(0<x<y<\infty)}}}$$

Calculate $P(X+Y \geq 1)$

Set where $A = \{(x,y): x+y \geq 1\}$ and $f_{xy}(x,y) \neq 0$.

$y \geq x$ and $x+y \geq 1$



A = Area bounded by lines:
$x = y$
$x+y = 1$
$x = 0$

Figure 4.1.1. *Regions for Example 4.1.12*

Graphs almost always needed!

Want: $P(X+Y \geq 1) \rightarrow$ integrate over purple region Set A

$\Rightarrow$ or calculate $1 - P(X+Y<1) = 1 -$ blue region Set B

"easier"

Know in Set B: $0<x<y<\infty$ and $x+y<1 \rightarrow y<1-x$

or $x<y$ and $y<1-x$

$x<y<1-x$

1st integrate over y

Region B bounded by
$x=0$ $0<x$
$x+y=1$, $x=y$ > max x is $\frac{1}{2}$

$$\therefore P(X+Y \geq 1) = 1 - P(X+Y<1) = 1 - \int_0^{1/2}\int_x^{1-x} e^{-y}\,dy\,dx$$

$$= 1 - \int_0^{1/2} e^{-x} - e^{-(1-x)}\,dx = 1 - \left[-e^{-x} + \frac{e^{-(1-x)}}{(-1)}\right]\Big|_0^{1/2}$$

$$= \cancel{1} - \left[(-e^{-1/2} - e^{-1/2}) - (\cancel{-1} - e^{-1})\right] = 2e^{-1/2} - e^{-1} \doteq .845$$

Have joint pdf (or pmf); what about joint cdf?

Joint cdf can also describe joint probability dist'n.

$$F(x,y) = P(X \le x, Y \le y) \quad \forall (x,y) \in \mathbb{R}^2$$

"joint cdf usually not very handy to use for discrete random vector".

But continuous

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)\, dt\, ds$$

$\Rightarrow$ implies

$$\frac{\partial^2 F(x,y)}{\partial x\, \partial y} = f(x,y)$$

$\Rightarrow$ useful when we can find expression for $F(x,y)$.

§4.2 Conditional Dist'ns and Independence

The values of (X,Y) our 2 Random Variables (X,Y) are related (dice example: knowing Y (|difference|) gives information about X (sum) and vice versa.

- X = height Y = weight; knowing X gives info about Y.
- Sometimes knowing X gives no info about Y.
- If (X,Y) discrete (countable) consider $P(Y=y \mid X=x)$, given $P(X=x)>0$

Definition 1.3.2 If A and B are events in S, and $P(B) > 0$, then the *conditional probability of A given B*, written $P(A|B)$, is

(1.3.1) $$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

$$P(Y=y \mid X=x) = \frac{P(Y=y \text{ and } X=x)}{P(X=x)} = \frac{P(Y=y, X=x)}{P(X=x)} \quad \text{for } P(X=x)>0$$

$$= f(x,y) / f_X(x)$$

Definition 4.2.1 Let (X, Y) be a discrete bivariate random vector with joint pmf $f(x,y)$ and marginal pmfs $f_X(x)$ and $f_Y(y)$. For any x such that $P(X=x) = f_X(x) > 0$, the *conditional pmf of* Y *given that* $X = x$ is the function of y denoted by $f(y|x)$ and defined by

$$f(y|x) = P(Y = y|X = x) = \frac{f(x,y)}{f_X(x)}.$$

For any y such that $P(Y = y) = f_Y(y) > 0$, the *conditional pmf of* X *given that* $Y = y$ is the function of x denoted by $f(x|y)$ and defined by

$$f(x|y) = P(X = x|Y = y) = \frac{f(x,y)}{f_Y(y)}.$$

Verify $f(y|x)$ is a pmf.

- since $f(x,y)$ + $f_X(x)$ are pmfs, $f(x,y) > 0$ + $f_X(x) > 0$.

$$\sum_y f(y|x) = \sum_y f(x,y) \Big/ f_X(x) = \frac{f_X(x)}{f_X(x)} = 1$$

Example: Calculating conditional Probs. of Y given X occurred.

Define $f(x,y)$: $f(0,10) = f(0,20) = \frac{2}{18}$ $\quad f(1,10) = f(1,30) = \frac{3}{18}$

$f(1,20) = \frac{4}{18}$ + $f(2,30) = 4/18$

Possible values X:

$X=0 \quad f_X(0) = \sum_Y f_{XY}(0,Y) = f(0,10) + f(0,20) = 4/18$

$X=1 \quad f_X(1) = \sum_Y f_{XY}(1,Y) = f(1,10) + f(1,30) + f(1,20) = 10/18$

$X=2 \quad f_X(2) = \sum_Y f_{XY}(2,Y) = f(2,30) = 4/18$

Note $X=0 \Rightarrow Y = 10$ or 20

Given $X=0 \Rightarrow f(y|X=0) > 0$ for $y = 10$ or 20

$$f_{Y|X}(10|X=0) = \frac{f(0,10)}{f_X(0)} = \frac{\frac{2}{18}}{\frac{4}{18}} = \frac{1}{2}; \qquad f_{Y|X}(20|X=0) = \frac{f(0,20)}{f_X(0)} = \frac{1}{2}$$

Similarly

$$f_{XY}(20|1) = f_{Y|X}(10|1) = \frac{f(1,10)}{f_X(1)} = \frac{\frac{3}{18}}{\frac{10}{18}} = \frac{3}{10}; \quad f_{Y|X}(20|1) = \frac{f(1,20)}{f_X(1)} = \frac{4/18}{10/18} = \frac{4}{10}$$

and

$$\underbrace{f_{XY}(30|2) = \frac{f(2,30)}{f_X(2)} = \frac{4/18}{4/18}}_{\text{If } X=2 \ Y \text{ must be } 30.} = 1.$$

Note: $\sum_Y f_{Y|X}(y|X=0) = 1$

$\sum_Y f_{Y|X}(y|X=1) = 1$

$\sum_Y f_{Y|X}(y|X=2) = 1$

Other probs:

$P(Y>10|X=1) = f(20|1) + f(30|1) = 7/10$ + $P(Y>10|X=0) = f(20|0) = \frac{1}{2}.$

Continuous X & Y

If X is continuous $P(X=x) = \int_x^x f_X(t)\,dt = 0$

$P(A|B) = \frac{P(A\cap B)}{P(B)}$ not applicable if $P(B)=0$.

Although, we will observe an X value, which may give info about Y.

Note that although $P(X=x)=0$ in continuous case $f_X(x)>0$ for $I_{(x\in S)}^{(x)}=1$.

Recall:

Definition 1.6.3 The *probability density function* or *pdf*, $f_X(x)$, of a continuous random variable X is the function that satisfies

(1.6.3) $$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt \quad \text{for all } x.$$

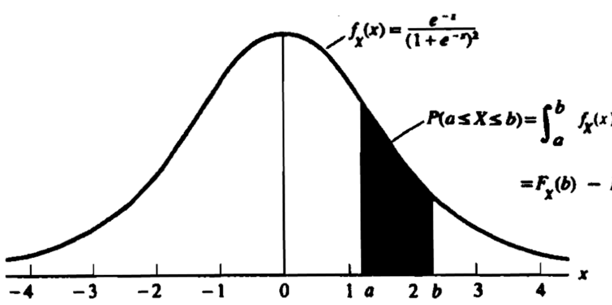


Figure 1.6.1. *Area under logistic curve*

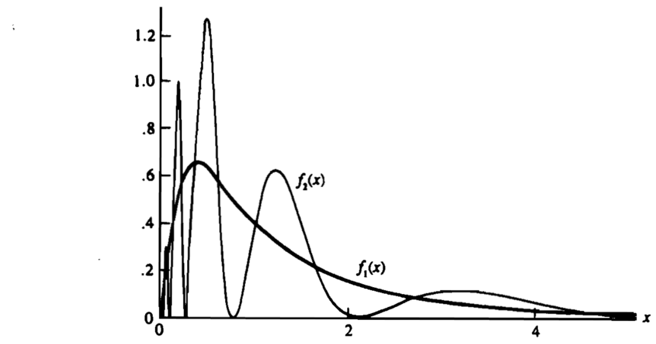


Figure 2.3.2. *Two pdfs with the same moments:* $f_1(x) = \frac{1}{\sqrt{2\pi}x}e^{-(\log x)^2/2}$ *and* $f_2(x) = f_1(x)[1+\sin(2\pi\log x)]$

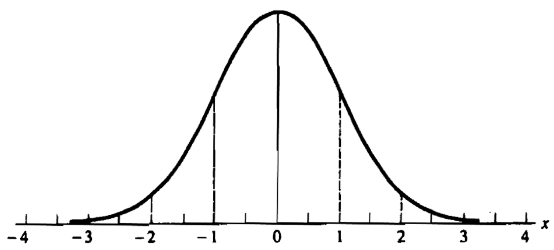


Figure 3.3.1. *Standard normal density*

→ The appropriate way to define $f_{Y|X}(y|x)$, analogous to discrete case, replace pmf by pdf (details beyond C+B).

Definition 4.2.3 Let (X, Y) be a continuous bivariate random vector with joint pdf $f(x,y)$ and marginal pdfs $f_X(x)$ and $f_Y(y)$. For any x such that $f_X(x) > 0$, the *conditional pdf of Y given that $X = x$* is the function of y denoted by $f(y|x)$ and defined by

$$f(y|x) = \frac{f(x,y)}{f_X(x)}.$$

For any y such that $f_Y(y) > 0$, the *conditional pdf of X given that $Y = y$* is the function of x denoted by $f(x|y)$ and defined by

$$f(x|y) = \frac{f(x,y)}{f_Y(y)}.$$

Verify pdf's (similar to discrete)

$f(x,y) > 0 + f_X(x) > 0 \Rightarrow f_{Y|X}(y|x) > 0$

$$\int_y f(y|x)\,dy = \int_y \frac{f(x,y)}{f_X(x)}\,dy = \frac{1}{f_X(x)} \int_y f(x,y)\,dy = \frac{f_X(x)}{f_X(x)} = 1$$

Example (Calculating conditional pdfs)

Recall example $f(x,y) = e^{-y} I_{(0<x<y<\infty)}(x,y)$.

Calculate $f(y|X=x)$.

If $x \le 0$ $I_{(0<x<y<\infty)}(x,y) = 0$; $f(x,y) = 0 \; \forall y$, so $f_X(x) = 0$.

If $x > 0$ $f(x,y) > 0$ if $y > x$:

notation abuse

So: $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_x^{\infty} e^{-y}\,dy = [-e^{-\infty} + e^{-x}] = e^{-x}$

$$\therefore \quad f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, \quad \text{if } y > x$$

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{0}{e^{-x}} = 0 \quad \text{if } y \le x$$

sample space depends on x, which we assume is given.

$$f(y|x) = e^{-(y-x)} I_{(x,\infty)}(y)$$

← Given $X = x$, $Y \sim$ exponential dist'n w/ location parameter x. (scale param $\beta = 1$).

Expectations for conditional dist'ns

$$E[g(Y)|x] = \sum_y g(y)\, f(y|x) \quad \leftarrow \text{discrete}$$

$$E[g(Y)|x] = \int_{-\infty}^{\infty} g(y)\, f(y|x)\,dy \quad \leftarrow \text{continuous}$$

For Exponential $f(y|x)$

Expected Value depends on x.

$$E[Y|X=x] = \int_x^{\infty} y e^{-(y-x)}\,dy = 1 + x.$$

Variance same all x

$$\mathrm{Var}(Y|x) = E[Y^2|x] - (E[Y|x])^2 = \int_x^{\infty} y^2 e^{-(y-x)}\,dy - \left(\int_x^{\infty} y e^{-(y-x)}\,dy\right)^2 = 1.$$