Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Dynamic Systems and Control
Institut für Dynamische Systeme und Regelungstechnik

Introduction to Recursive Filtering and Estimation

Spring 2010

Problem Set: Probability review

Notes:

- **Notation:** Unless otherwise noted, x, y, and z denote random variables, $f_x(x)$ (or the short hand $f(x)$) denotes the probability density function of x, and $f_{x|y}(x|y)$ (or $f(x|y)$) denotes the conditional probability density function of x conditioned on y. The expected value is denoted by $\mathrm{E}[\cdot]$, the variance is denoted by $\mathrm{Var}(\cdot)$ and $\Pr(Z)$ denotes the probability that the event Z occurs.
- Please report any error that you may find to the teaching assistants (`strimpe@ethz.ch` or `aschoellig@ethz.ch`).

Problem Set

Problem 1

Prove $f(x|y) = f(x) \Leftrightarrow f(y|x) = f(y)$.

Problem 2

Prove $f(x|y,z) = f(x|z) \Leftrightarrow f(x,y|z) = f(x|z)f(y|z)$.

Problem 3

Come up with an example where $f(x,y|z) = f(x|z)f(y|z)$ but $f(x,y) \neq f(x)f(y)$ for *continuous* random variables x, y, and z.

Problem 4

Come up with an example where $f(x,y|z) = f(x|z)f(y|z)$ but $f(x,y) \neq f(x)f(y)$ for *discrete* random variables x, y, and z.

Problem 5

Let x and y be binary random variables: $x, y \in \{0,1\}$. Prove that $f_{x,y}(x,y) \geq f_x(x) + f_y(y) - 1$, which is called *Bonferroni's inequality*.

Problem 6

Prove $\mathrm{E}[ax+b] = a\,\mathrm{E}[x] + b$, where a and b are constants.

Problem 7

Let x and y be scalar random variables. Prove that for any functions $g(\cdot)$ and $h(\cdot)$, $\mathrm{E}[g(x)h(y)] = \mathrm{E}[g(x)]\,\mathrm{E}[h(y)]$ if x and y are independent.

Problem 8

From above, it follows that if x and y are independent, then $\mathrm{E}[xy] = \mathrm{E}[x]\,\mathrm{E}[y]$. Is the converse true, that is, if $\mathrm{E}[xy] = \mathrm{E}[x]\,\mathrm{E}[y]$, does it imply that x and y are independent? If yes, prove it. If no, find a counter-example.

Problem 9

Let $x \in \mathcal{X}$ be a scalar valued continuous random variable and $g(x)$ a continuously differentiable, strictly monotonically increasing function. Prove the *law of the unconscious statistician*, that is, show that the expected value of $y = g(x)$ is given by

$$\mathrm{E}[y] = \int\limits_{x \in \mathcal{X}} g(x) f_x(x)\, dx.$$

Problem 10

Let y be a scalar continuous random variable with probability density function $f_y(y)$, and g be a function transforming y to a new variable x by $x = g(y)$. Prove the *change of variables formula*, i.e. the equation for the probability density function $f_x(x)$, under the assumption that g is continuously differentiable with $\frac{dg(y)}{dy} < 0 \; \forall y$.

Problem 11

Let $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ be independent, scalar valued discrete random variables. Let $z = x + y$. Prove that

$$f_z(z) = \sum_{y \in \mathcal{Y}} f_x(z - y) f_y(y) = \sum_{x \in \mathcal{X}} f_x(x) f_y(z - x),$$

thus the probability density function is the *convolution* of the individual probability density functions.

Problem 12

Let x be a scalar continuous random variable that takes on only non-negative values. Prove that, for $a > 0$,

$$\Pr(x \geq a) \leq \frac{\mathrm{E}[x]}{a},$$

which is called *Markov's inequality.*

Problem 13

Let x be a scalar continuous random variable. Let $\bar{x} = \mathrm{E}[x]$, $\sigma^2 = \mathrm{Var}(x)$. Prove that for any $k > 0$,

$$\Pr\left(|x - \bar{x}| \geq k\right) \leq \frac{\sigma^2}{k^2},$$

which is called *Chebyshev's inequality.*

Problem 14

Let $x_1, x_2, \ldots, x_n$ be independent random variables, each having a uniform distribution over $[0, 1]$. Let $m = \max\{x_1, x_2, \ldots, x_n\}$. Show that the cumulative distribution function of m, $F_m(\cdot)$, is given by

$$F_m(m) = m^n, \quad 0 \leq m \leq 1.$$

What is the probability density function of m?

Problem 15

Prove that $\mathrm{E}[x^2] \geq (\mathrm{E}[x])^2$. When do we have equality?

Problem 16

Suppose that x and y are independent scalar continuous random variables. Show that

$$\Pr(x \leq y) = \int\limits_{-\infty}^{\infty} F_x(y) f_y(y)\, dy.$$

Problem 17

Use Chebyshev's inequality to prove the *weak law of large numbers*. Namely, if $x_1, x_2, \dots$ are independent and identically distributed with mean μ and variance σ^2 then, for any $\varepsilon > 0$,

$$\Pr\left(\left|\frac{x_1 + x_2 + \cdots + x_n}{n} - \mu\right| > \varepsilon\right) \to 0 \quad \text{as} \quad n \to \infty.$$

Problem 18

Suppose that x is a random variable with mean 10 and variance 15. What can we say about $\Pr(5 < x < 15)$?

Problem 19

Let x and y be independent random variables with means μ_x and μ_y and variances σ_x^2 and σ_y^2. Show that

$$\mathrm{Var}(xy) = \sigma_x^2 \sigma_y^2 + \mu_y^2 \sigma_x^2 + \mu_x^2 \sigma_y^2.$$

Problem 20

a) Use the method presented in class to obtain samples x of the *exponential distribution*

$$\hat{f}_x(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

from given samples u of a uniform distribution.

b) Exponential random variables are used to model failures of components that do not wear, for example solid state electronics. This is based on their property

$$\Pr(x > s + t | x > t) = \Pr(x > s) \;\; \forall s, t \geq 0.$$

A random variable with this property is said to be *memoryless*. Prove that a random variable with an exponential distribution satisfies this property.

Problem Set 1: Recursive Filtering & Estimation

Probability Review

Problem 1

- from definition of conditional probability, we have

$$f(x,y) = f(x|y)\,f(y) \quad \text{and} \quad f(x,y) = f(y|x)\,f(x)$$

and therefore $\quad f(x|y)\,f(y) = f(y|x)\,f(x) \qquad (*)$

- If $f(x|y) = f(x) \overset{(*)}{\Rightarrow} f(x)\,f(y) = f(y|x)\,f(x)$

$\Rightarrow f(y) = f(y|x)$

If $f(y|x) = f(y) \overset{(*)}{\Rightarrow} f(x|y)\,f(y) = f(y)\,f(x)$

$\Rightarrow f(x) = f(x|y)$

$\therefore \quad f(x|y) = f(x) \Leftrightarrow f(y|x) = f(y)$ q.e.d.

Problem 2

Follows directly from $f(x,y|z) = f(x|y,z)\,f(y|z)$

Problem 3

- Let $x, y, z \in [0,1]$ be continuous random variables.

- Consider joint pdf $f(x,y,z) = c \cdot g_x(x,z)\, g_y(y,z)$ where g_x and g_y are functions (not necessarily pdf's) and c is a constant such that

$$\int_0^1\int_0^1\int_0^1 f(x,y,z)\, dx\, dy\, dz = 1$$

Then

$$f(x,y|z) = \frac{f(x,y,z)}{f(z)} = \frac{c \cdot g_x(x,z)\, g_y(y,z)}{c \cdot G_x(z) \cdot G_y(z)},$$

with

$$G_x(z) = \int_0^1 g_x(x,z)\, dx, \quad G_y(z) = \int_0^1 g_y(y,z)\, dy$$

(since $f(z) = \int_0^1\int_0^1 c\, g_x(x,z)\, g_y(y,z)\, dx\, dy$

$$= c\int_0^1 G_x(z) \cdot g_y(y,z)\, dy = c\, G_x(z) \cdot G_y(z).)$$

Also

$$f(x|z) = \frac{f(x,z)}{f(z)} = \frac{\int_0^1 f(x,y,z)\, dy}{f(z)} = \frac{c \cdot g_x(x,z)\, G_y(z)}{c \cdot G_x(z)\, G_y(z)}$$

$$= \frac{g_x(x,z)}{G_x(z)}$$

$$f(y|z) = \frac{g_y(y,z)}{G_y(z)} \quad \text{(by analogy)}$$

Therefore, $f(x,y|z) = \dfrac{g_x(x,z) \cdot g_y(y,z)}{G_x(z)\, G_y(z)} = f(x|z) \cdot f(y|z)$.

- Now, let $g_x(x,z) = x + z$ and $g_y(y,z) = y + z$.

Then

$$f(x,y) = \int_0^1 f(x,y,z)\, dz = c\int_0^1 (x+z)(y+z)\, dz$$

$$= c\int_0^1 z^2 + (x+y)z + xy\, dz = c\left[\tfrac{1}{3}z^3 + \tfrac{1}{2}(x+y)z^2 + xyz\right]_{z=0}^{z=1}$$

$$= c\left(\tfrac{1}{3} + \tfrac{1}{2}(x+y) + xy\right)$$

$$f(x) = \int_0^1 f(x,y)\,dx = c\int_0^1 \tfrac{1}{3} + \tfrac{1}{2}x + \tfrac{1}{2}y + xy\,dx$$

$$= c\left(\tfrac{1}{3} + \tfrac{1}{4} + \tfrac{1}{2}y + \tfrac{1}{2}y\right) = c\left(y + \tfrac{7}{12}\right)$$

$$f(y) = \int_0^1 f(x,y)\,dy = c\left(x + \tfrac{7}{12}\right) \qquad \text{(by symmetry)}$$

Therfore, $f(x)\,f(y) = c^2\left(\frac{49}{144} + \frac{7}{12}(x+y) + xy\right) \neq f(x,y)$.

Problem 4

- As in problem 3, let $f(x,y,z) = c\, g_x(x,z)\, g_y(y,z)$. This ensures $f(x,y\mid z) = f(x\mid z)\, f(y\mid z)$. Let $x, y, z \in \{0,1\}$

- Define g_x and g_y by value tables:

x	z	$g_x(x,z)$
0	0	0
0	1	1
1	0	1
1	1	0

y	z	$g_y(y,z)$
0	0	1
0	1	0
1	0	0
1	1	1

x	y	z	$f(x,y,z) = \frac{1}{2} g_x(x,z)\, g_y(y,z)$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0.5
1	0	0	0.5
1	0	1	0
1	1	0	0
1	1	1	0

• Compute $f(x,y)$, $f(x)$, $f(y)$

x	y	$f(x,y)$
0	0	0
0	1	0.5
1	0	0.5
1	1	0

x	$f(x)$
0	0.5
1	0.5

y	$f(y)$
0	0.5
1	0.5

x	y	$f(x) \cdot f(y)$
0	0	0.25
0	1	0.25
1	0	0.25
1	1	0.25

$$\left(\begin{array}{l} f(x,y) = \sum_z f(x,y,z) = f(x,y,0) + f(x,y,1) \\ f(x) = f(x,0) + f(x,1) \end{array} \right)$$

$\therefore \quad f(x,y) \neq f(x)\, f(y)$

• As verification, compute $f(x,y|z)$, $f(x|z)$, $f(y|z)$, although by construction of $f(x,y,z)$ it is already clear that $f(x,y|z) = f(x|z)\, f(y|z)$.

x	y	z	$f(x,y\|z)$
0	0	0	0
0	1	0	0
1	0	0	1
1	1	0	0
0	0	1	0
0	1	1	1
1	0	1	0
1	1	1	0

From $f(x,y,z)$ it follows that if $z=0$, then $(x,y)=(1,0)$ with probability 1

(or: compute from $f(x,y|z) = \dfrac{f(x,y,z)}{f(z)}$

similarly

x	z	$f(x\|z)$
0	0	0
1	0	1
0	1	1
1	1	0

y	z	$f(y\|z)$
0	0	1
1	0	0
0	1	0
1	1	1

x	y	z	$f(x\|z)\, f(y\|z)$
0	0	0	0
0	1	0	0
1	0	0	1
1	1	0	0
0	0	1	0
0	1	1	1
1	0	1	0
1	1	1	0

Problem 5

- By marginalization,

$$f_X(x) = \sum_{y \in Y} f_{XY}(x,y) = f_{XY}(x,y) + f_{XY}(x,1-y)$$

$$f_Y(y) = \sum_{x \in X} f_{XY}(x,y) = f_{XY}(x,y) + f_{XY}(1-x,y)$$

- Therefore,

$$f_X(x) + f_Y(y) = 2 f_{XY}(x,y) + f_{XY}(x,1-y) + f_{XY}(1-x,y)$$

$$= \underbrace{f_{XY}(x,y) + f_{XY}(x,1-y) + f_{XY}(1-x,y) + f_{XY}(1-x,1-y)}_{= \sum_{y \in \{0,1\}} \sum_{x \in \{0,1\}} f_{XY}(x,y) = 1}$$

$$- f_{XY}(1-x,1-y) + f_{XY}(x,y)$$

$$= 1 + f_{XY}(x,y) - f_{XY}(1-x,1-y)$$

$$\leq 1 + f_{XY}(x,y)$$

$$\Rightarrow \quad f_{XY}(x,y) \geq f_X(x) + f_Y(y) - 1$$

q.e.d.

Problem 6

Using the law of the unconscious statistician, we have for a discrete random variable X

$$E[aX+b] = \sum_{x \in X} (ax+b) f_X(x)\, dx = a \underbrace{\sum_{x \in X} x f_X(x)}_{E[X]} + b \underbrace{\sum_{x \in X} f_X(x)}_{1}$$

$$= a\,E[X] + b$$

and for a continuous random variable X

$$E[aX+b] = \int_{-\infty}^{\infty} (ax+b) f_X(x)\, dx$$

$$= a \underbrace{\int_{-\infty}^{\infty} x f_X(x)\, dx}_{E[X]} + b \underbrace{\int_{-\infty}^{\infty} f_X(x)\, dx}_{1}$$

$$= a\,E[X] + b$$

q.e.d.

Problem 7

Using the law of the unconscious statistician applied on the vector random variable $\begin{bmatrix} X \\ Y \end{bmatrix}$ with pdf $f_{XY}(x,y)$, we have

$$E[g(x)h(y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)\, f_{XY}(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)\, f_X(x)\, f_Y(y)\, dx\, dy$$

(by independence x, y)

$$= \int_{-\infty}^{\infty} g(x) f_X(x)\, dx \cdot \int_{-\infty}^{\infty} h(y) \cdot f_Y(y)\, dy$$

$$= E[g(x)] \cdot E[h(y)].$$

and similarly for discrete random variables

Problem 8

We will construct a counterexample.

- Let x be a discrete r.v. that takes the values -1, 0, and 1 with probabilities $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$, respectively.
- Let y be a discrete r.v. taking the values -1, 1 with equal probability. Let x and y be independent.
- Let z be a discrete r.v. defined by $z = x \cdot y$.(*) By construction, it is clear that z depends on x. We will show that $E[xz] = E[x] \cdot E[z]$.

(*) It follows that z can take the values $-1, 0, 1$.

Pdf $f_{xz}(x,z)$:

x	z	$f_{xz}(x,z)$	
-1	-1	1/8	$\rightarrow$ $x=-1$ (Prob $\frac{1}{4}$) $\wedge$ $w=1$ (Prob $\frac{1}{2}$)
-1	0	0	$\rightarrow$ impossible since $w \neq 0$
-1	1	1/8	$\rightarrow$ $x=-1$ (Prob $\frac{1}{4}$) $\wedge$ $w=-1$ (Prob $\frac{1}{2}$)
0	-1	0	$\rightarrow$ impossible
0	0	1/2	$\rightarrow$ w either ± 1, $x=0$ (Prob. $\frac{1}{2}$)
0	1	0	⋮
1	-1	1/8	
1	0	0	
1	1	1/8	

Marginalization: $f_x(x)$, $f_z(z)$

x	$f(x)$
-1	1/4
0	1/2
1	1/4 ✓

z	$f(z)$
-1	1/4
0	1/2
1	1/4 ✓

It follows that $f_{xz}(x,z) = f_x(x)\, f_z(z) \ \forall x,z$ does not hold.
For example, $f_{xz}(-1,0) = 0 \neq 1/8 = f_x(-1)\, f_z(0)$

Expected values:

$$E[xz] = \sum_x \sum_z xz\, f_{xz}(x,z)$$
$$= (-1)(-1)\tfrac{1}{8} + (-1)(1)\tfrac{1}{8} + 0 \cdot 0 \cdot \tfrac{1}{2} + (1)(-1)\cdot\tfrac{1}{8} + (1)(1)\tfrac{1}{8}$$
$$= 0$$
$$E[x] = \sum_x x\, f_x(x)$$
$$= -\tfrac{1}{4} + \tfrac{1}{4} = 0$$
$$E[z] = 0$$

$$\therefore\ E[xz] = E[x]\,E[z] \ \not\Rightarrow\ f_{xz}(x,z) = f_x(x)\cdot f_z(z)$$

Problem 9

- From the change of variables formula (see problem 10), we have

$$f_Y(y) = \frac{f_X(x)}{\frac{dg}{dx}(x)} \quad \text{(since } g \text{ cont differentiable, strictly monotonic)}$$

- Furthermore, $y = g(x) \Rightarrow dy = \frac{dg}{dx}(x)\,dx$

- Let $\mathbb{Y} = g(\mathbb{X})$, that is $\mathbb{Y} = \{ y \mid \exists x \in \mathbb{X} : y = g(x) \}$.

- $$E[y] = \int_{y \in \mathbb{Y}} y\, f_Y(y)\, dy = \int_{x \in \mathbb{X}} g(x) \frac{f_X(x)}{\cancel{\frac{dg}{dx}(x)}} \cancel{\frac{dg}{dx}(x)}\, dx$$

$$= \int_{x \in \mathbb{X}} g(x)\, f_X(x)\, dx$$

q.e.d.

Problem 10

- Consider the probability that y is in a small interval $[\bar{y}, \bar{y} + \Delta y]$ (with $\Delta y > 0$ small),

$$P(y \in [\bar{y}, \bar{y} + \Delta y]) = \int_{\bar{y}}^{\bar{y}+\Delta y} f_Y(y)\, dy$$

which approaches $f_Y(\bar{y})\Delta y$ in the limit as $\Delta y \to 0$

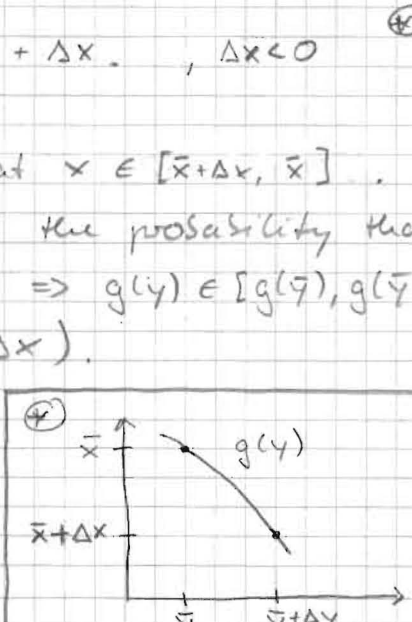
- Let $\bar{x} := g(\bar{y})$. For small Δy, we have (Taylor series)

$$g(\bar{y} + \Delta y) \approx g(\bar{y}) + \underbrace{\frac{dg}{dy}(\bar{y}) \cdot \Delta y}_{=: \Delta x} = \bar{x} + \Delta x. \quad , \Delta x < 0 \quad ⊛$$

- We are interested in the probability that $x \in [\bar{x} + \Delta x, \bar{x}]$.
For small Δx, this probability is equal the probability that $y \in [\bar{y}, \bar{y} + \Delta y]$ (since $y \in [\bar{y}, \bar{y} + \Delta y] \Rightarrow g(y) \in [g(\bar{y}), g(\bar{y} + \Delta y)]$ $\Rightarrow x \in [\bar{x} + \Delta x, \bar{x}]$ for small Δy and Δx).

$$\therefore \; P(y \in [\bar{y}, \bar{y} + \Delta y]) = f_Y(\bar{y})\Delta y$$
$$= P(x \in [\bar{x} + \Delta x, \bar{x}]) = -f_X(\bar{x})\Delta x$$

for small Δy (and thus small Δx)



g(y) decreasing ⇒ Δx neg.

$$\Rightarrow \quad f_Y(\bar{y})\Delta y = -f_X(\bar{x})\Delta x$$

$$\Rightarrow \quad f_X(\bar{x}) = \frac{f_Y(\bar{y})}{-\frac{\Delta x}{\Delta y}}$$

- As $\Delta y \to 0$ (and thus $\Delta x \to 0$), $\frac{\Delta x}{\Delta y} \to + \frac{dg}{dy}(\bar{y})$ and therefore

$$f_X(\bar{x}) = \frac{f_Y(\bar{y})}{-\frac{dg}{dy}(\bar{y})}$$

or

$$f_X(x) = \frac{f_Y(y)}{-\frac{dg}{dy}(y)} \quad .$$

Problem 11

- $f_Z(\bar{z}) = Pr\{z = \bar{z}\}$

The probability that z takes on $\bar{z}$ is given by the sum of the probabilities of all possible combinations of x and y such that $\bar{z} = x + y$, that is all $\bar{y} \in Y$ s.t. $x = \bar{z} - \bar{y}$ and $y = \bar{y}$ [or alternatively, $\bar{x} \in X$ s.t. $x = \bar{x}$ and $y = \bar{z} - \bar{x}$].

$$\therefore \quad f_Z(\bar{z}) = Pr\{z = \bar{z}\} = \sum_{\bar{y} \in Y} Pr\{x = \bar{z} - \bar{y}\} \cdot Pr\{y = \bar{y}\} \quad \text{(by independence assumption)}$$

$$= \sum_{\bar{y} \in Y} f_X(\bar{z} - \bar{y}) f_Y(\bar{y})$$

Rewrite $f_Z(z) = \sum_{y \in Y} f_X(z-y) f_Y(y)$

Similarly, $f_Z(z) = \sum_{x \in X} f_X(x) f_Y(z-x)$.

q.e.d.

Problem 12

$$E[x] = \int_0^\infty x\, f(x)\, dx = \underbrace{\int_0^a x\, f(x)\, dx}_{\geq 0} + \int_a^\infty x\, f(x)\, dx$$

$$\geq \int_a^\infty x\, f(x)\, dx \geq \int_a^\infty a\, f(x)\, dx$$

$$= a \int_a^\infty f(x)\, dx = a \Pr(x \geq a)$$

q.e.d.

Problem 13

Can be obtained as a corollary of Markov's inequality:

- $(x-\bar{x})^2$ is a nonnegative random variable; apply Markov's inequality with $a = k^2$ $(k>0)$:

$$\Pr\left((x-\bar{x})^2 \geq k^2\right) \leq \frac{E[(x-\bar{x})^2]}{k^2} = \frac{\mathrm{Var}(x)}{k^2} = \frac{\sigma^2}{k^2}$$

- Since $(x-\bar{x})^2 \geq k^2 \Leftrightarrow |x-\bar{x}| \geq k$,

$$\Pr(|x-\bar{x}| \geq k) \leq \frac{\sigma^2}{k^2}$$

q.e.d.

Problem 14

For $0 \leq t \leq 1$,

$$F_m(t) = \Pr(m \in [0,t]) = \Pr(\max\{x_1, \ldots, x_n\} \in [0,t])$$

$$= \Pr(x_i \in [0,t]\ \forall i \in \{1,\ldots,n\})$$

$$= \Pr(x_1 \in [0,t]) \cdot \Pr(x_2 \in [0,t]) \cdot \ldots \cdot \Pr(x_n \in [0,t])$$

$$= t^n \qquad (\text{with } t = m \to F_m(m) = m^n)$$

Since $F_m(m) = \int_0^m f_m(\tau)\, d\tau$,

$$f_m(m) = \frac{dF_m}{dm}(m) = \frac{d}{dm}(m^n) = n \cdot m^{n-1}.$$

Problem 15

Let $\bar{x} = E[x]$

$$Var(x) = \int_{-\infty}^{\infty} (x-\bar{x})^2 f(x)\,dx = \int_{-\infty}^{\infty} x^2 f(x)\,dx - 2\bar{x} \underbrace{\int_{-\infty}^{\infty} x f(x)\,dx}_{\bar{x}} + \bar{x}^2 \underbrace{\int_{-\infty}^{\infty} f(x)\,dx}_{1}$$

$$= E[x^2] - 2\bar{x}^2 + \bar{x}^2 = E[x^2] - \bar{x}^2$$

$$\therefore \quad E[x^2] = \bar{x}^2 + Var(x)$$
$$\geq \bar{x}^2 \qquad (\text{since } Var(x) \geq 0)$$

with equality if $Var(x) = 0$, i.e. when x is constant.

Problem 16

- For fixed $\bar{y} \in Y$,

  $Pr(x \leq \bar{y}) = F_x(\bar{y})$

- The probability that $y \in [\bar{y}, \bar{y}+\Delta y]$ is

  $Pr(y \in [\bar{y}, \bar{y}+\Delta y]) = \int_{\bar{y}}^{\bar{y}+\Delta y} f_y(y)\,dy \rightarrow f_y(\bar{y}) \cdot \Delta y$ for small Δy

- Therefore, the probability that $x \leq \bar{y}$ and $y \in [\bar{y}, \bar{y}+\Delta y]$ is

  $Pr(x \leq \bar{y} \wedge y \in [\bar{y}, \bar{y}+\Delta y]) = Pr(x \leq \bar{y}) \cdot Pr(y \in [\bar{y}, \bar{y}+\Delta y])$

  $= F_x(\bar{y}) \cdot f_y(\bar{y}) \Delta y \quad (*) \qquad$ (by indep. of x, y)

- Now, since we are interested in $Pr(x \leq y)$, i.e. in the probability of x being less than <u>any</u> y and not just a fixed $\bar{y}$, we can sum $(*)$ over all possible intervals $[\bar{y}_i, \bar{y}_i + \Delta y]$

  $$Pr(\exists \bar{y}_i \text{ such that } x \leq \bar{y}_i) = \sum_{i=-\infty}^{\infty} F_x(\bar{y}_i) \cdot f_y(\bar{y}_i) \cdot \Delta y$$

  By letting $\Delta y \rightarrow 0$, we obtain

  $$Pr(x \leq y) = \int_{-\infty}^{\infty} F_x(y)\, f_y(y)\,dy .$$

Problem 17

- Consider the random variable $\frac{X_1 + X_2 + \ldots + X_n}{n}$.

It has the mean

$$E\left[\frac{X_1 + X_2 + \ldots + X_n}{n}\right] = \frac{1}{n}\left(E[X_1] + \ldots + E[X_n]\right) = \frac{1}{n}(n \cdot \mu) = \mu$$

and variance

$$\operatorname{Var}\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = E\left[\left(\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right)^2\right]$$

$$= \frac{1}{n^2}\left(E\left[(X_1-\mu)^2 + \ldots + (X_n-\mu)^2 + \text{"cross terms"}\right]\right)$$

where the expected value of the cross terms is zero, for example,

$$E[(X_i-\mu)(X_j-\mu)] = E[X_i-\mu] \cdot E[X_j-\mu] \text{ for } i \neq j \text{, by independence}$$
$$= 0$$

$$\therefore \operatorname{Var}\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{1}{n^2}\left(\operatorname{Var}(X_1) + \ldots + \operatorname{Var}(X_n)\right) = \frac{n \cdot \sigma^2}{n^2} = \frac{\sigma^2}{n}$$

- Applying Chebyshev's inequality with $\epsilon > 0$:

$$\Pr\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n \cdot \epsilon^2} \rightarrow 0 \text{ as } n \to \infty$$

$$\Rightarrow \Pr\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) \to 0 \text{ as } n \to \infty$$

Note: $\Pr\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| \geq \epsilon\right) =$

$$\Pr\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) + \underbrace{\Pr\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| = \epsilon\right)}_{=0 \text{ since } \frac{X_1 + \ldots + X_n}{n} \text{ is a continuous random variable}}$$

Problem 18

- $\Pr(5 < x < 15)$
  $= \Pr(|x-10| < 5) = 1 - \Pr(|x-10| \geq 5)$

- Chebyshev's inequality with $\bar{x} = 10$, $\sigma^2 = 15$, $k = 5$
  $\Pr(|x-10| \geq 5) \leq \frac{15}{5^2} = \frac{15}{25} = \frac{3}{5}$

$\therefore \quad \Pr(5 < x < 15) \geq 1 - \frac{3}{5} = \frac{2}{5}$.

Problem 19

- Mean: $E[xy] = E[x] \cdot E[y] = \mu_x \mu_y$

- $\mathrm{Var}(xy) = E[(xy)^2] - (\mu_x \mu_y)^2$
  $= E[x^2]\, E[y^2] - \mu_x^2 \mu_y^2$

- Using $E[x^2] = \sigma_x^2 + \mu_x^2$ and $E[y^2] = \sigma_y^2 + \mu_y^2$, we have

  $\mathrm{Var}(xy) = (\sigma_x^2 + \mu_x^2)(\sigma_y^2 + \mu_y^2) - \mu_x^2 \mu_y^2$
  $= \sigma_x^2 \sigma_y^2 + \mu_y^2 \sigma_x^2 + \mu_x^2 \sigma_y^2$

q.e.d.

Problem 20

a) • cummulative distribution:

$$\hat{F}_X(x) = \int_0^x \lambda e^{-\lambda t}\, dt = \left[-e^{-\lambda t}\right]_{t=0}^{x} = 1 - e^{-\lambda x}$$

• u: sample from uniform distribution. According to method presented in class, obtain sample x by solving $u = \hat{F}_X(x)$ for x:

$$u = 1 - e^{-\lambda x} \quad \Leftrightarrow \quad e^{-\lambda x} = 1-u$$

$$\Leftrightarrow \quad -\lambda x = \ln(1-u)$$

$$\Leftrightarrow \quad x = \frac{-\ln(1-u)}{\lambda}$$

• $u \to 0 \Rightarrow x \to 0$,

$u \to 1 \Rightarrow x \to +\infty$

b)

$$Pr(x > s+t \mid x > t)$$

$$= \frac{Pr(x > s+t \wedge x > t)}{Pr(x>t)} = \frac{Pr(x > s+t)}{Pr(x > t)}$$

$$= \frac{\int_{s+t}^{\infty} \lambda e^{-\lambda x}\, dx}{\int_{t}^{\infty} \lambda e^{-\lambda x}\, dx} = \frac{\left[-e^{-\lambda x}\right]_{s+t}^{\infty}}{\left[-e^{-\lambda x}\right]_{t}^{\infty}} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}}$$

$$= e^{-\lambda s} = \int_s^{\infty} \lambda e^{-\lambda x}\, dx$$

$$= Pr(x > s)$$

q.e.d.