



Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación

Clase 24

Métodos estadísticos en ingeniería de software

IIC2143 - Ingeniería de Software
Sección 1

Rodrigo Saffie

rasaffie@uc.cl

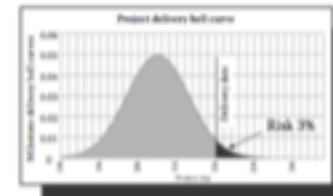
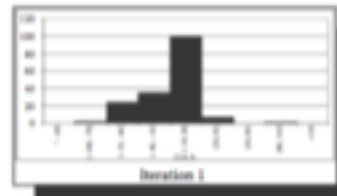
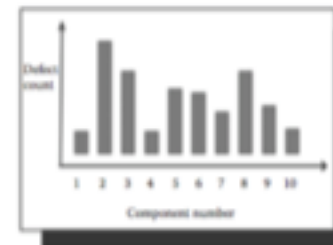
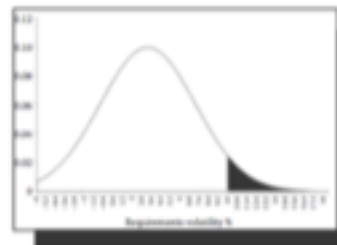
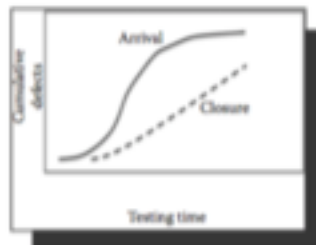
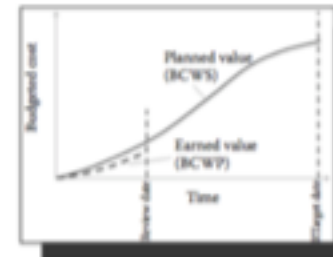
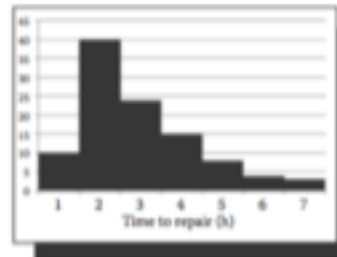
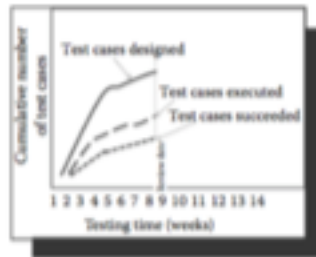
20 de junio de 2018

Etapas de una métrica

- **Formulación:** formalización de factores apropiados para representar el *software*
- **Recolección:** mecanismos para acumular datos a partir de la formulación
- **Análisis:** procesamiento de los valores recolectados para obtener información
- **Interpretación:** evaluación de la información para determinar mejoras
- **Retroalimentación:** recomendaciones para el equipo de desarrollo, derivadas de la interpretación

Visualización de métricas

- Las métricas son representadas con gráficos



Características de una buena métrica

- Simple y computable
- Intuitiva
- Consistente y objetiva
- Unidades de medición expresivas
- Reflejar recomendaciones para mejorar
- Independiente del contexto (equipo, lenguaje de programación)

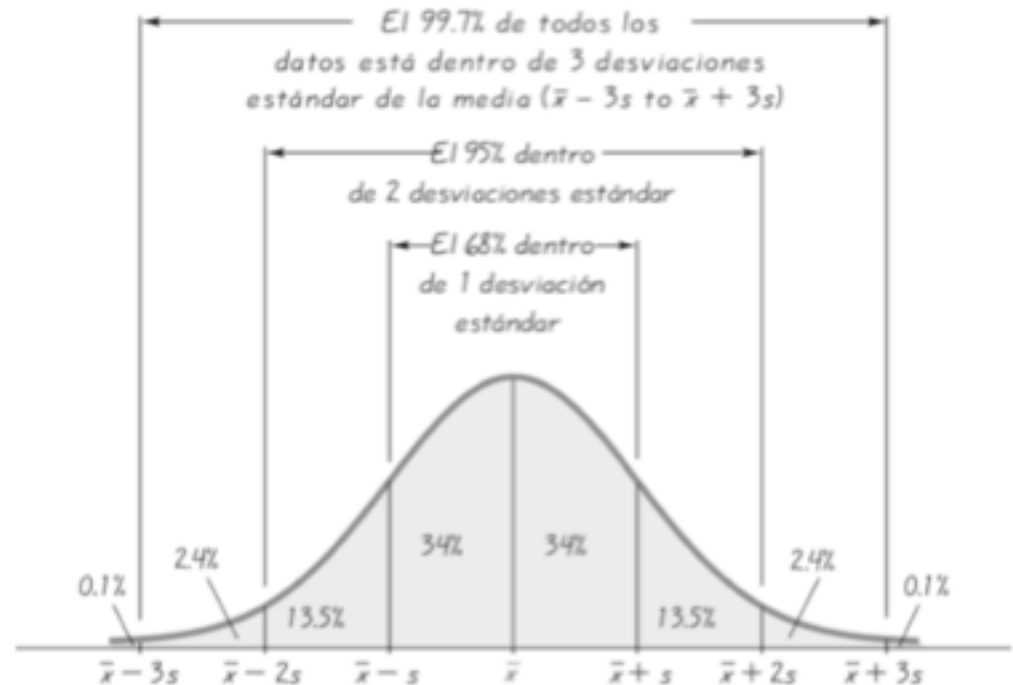
Ejemplos de métricas

- Métricas de *testing*
- Métricas de diseño
- Métricas aplicaciones web

Métricas aplicaciones web

- **Tiempos de respuesta**

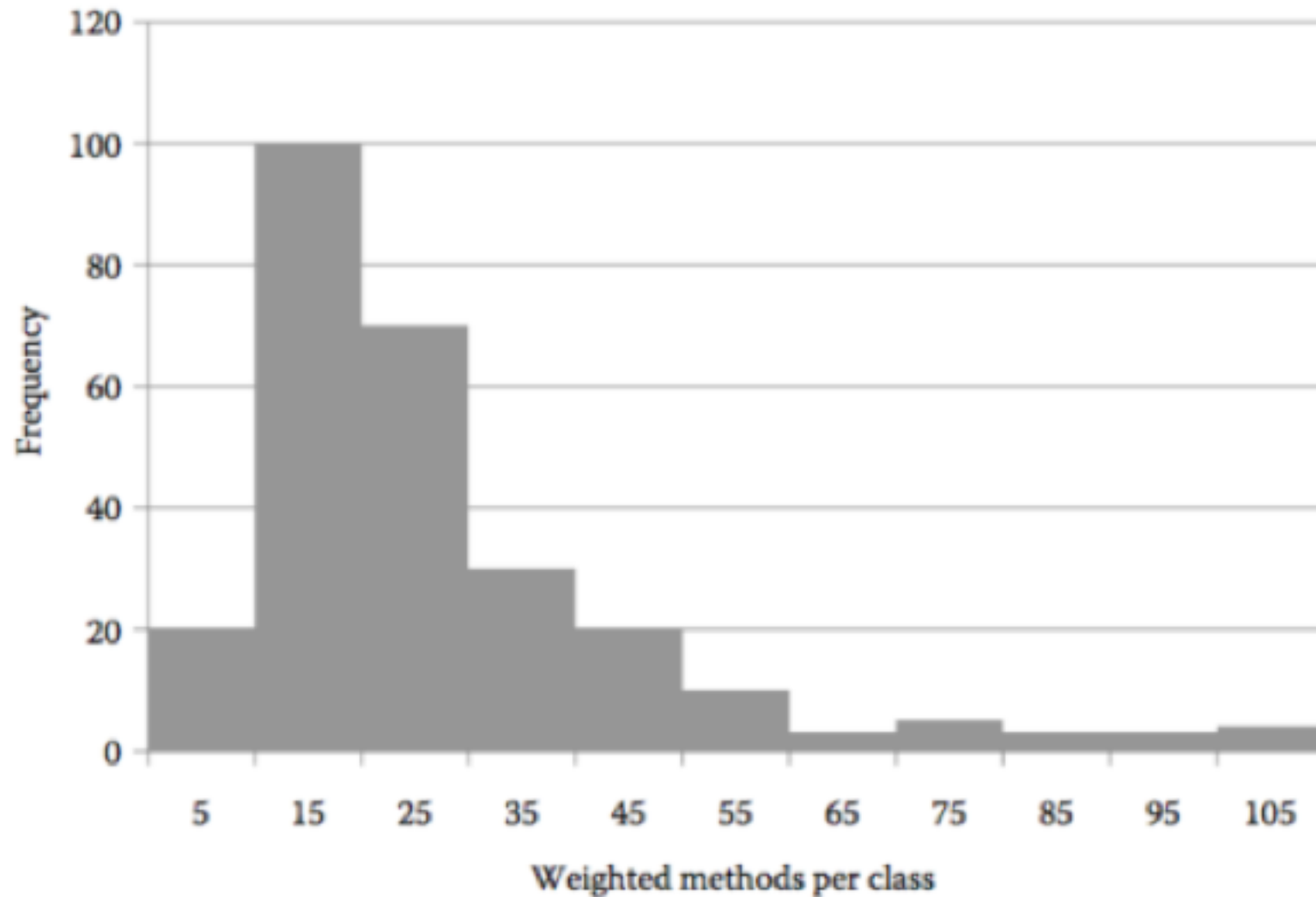
- Media
- Mediana
- Moda
- Desviación estándar



Métricas aplicaciones web

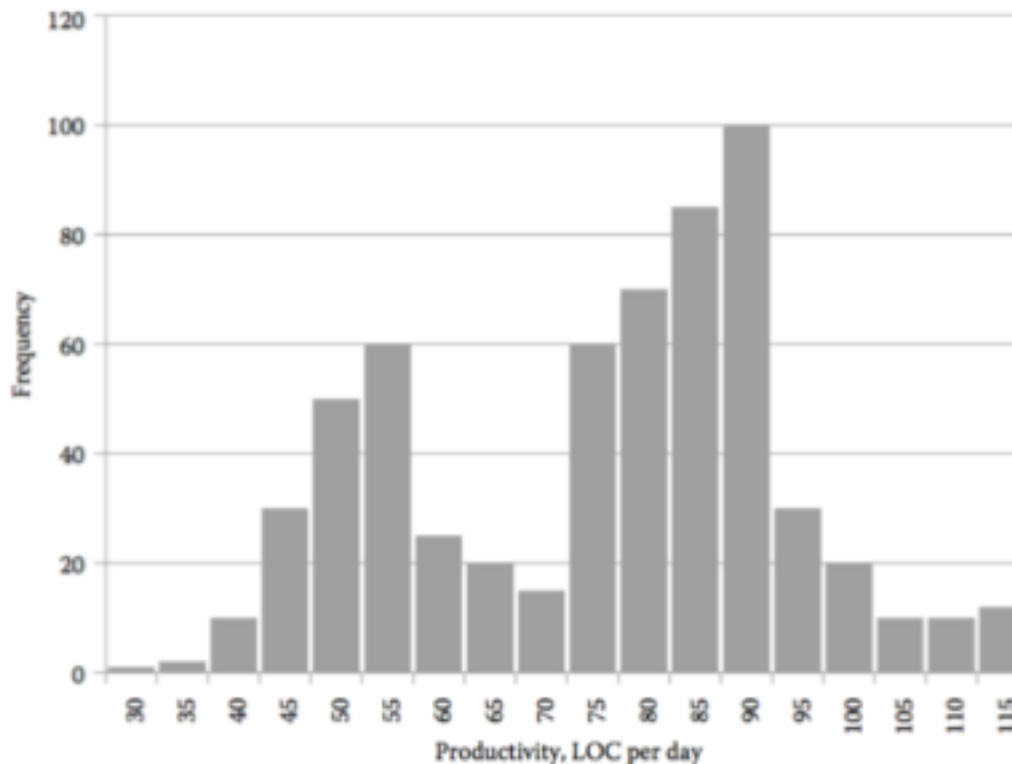
- **Tiempos de respuesta**
 - Media
 - Mediana
 - Moda
 - Desviación estándar
- **Throughput:** cantidad de solicitudes por minuto

Histogramas



Forma del histograma

- Cambia radicalmente según intervalos escogidos
- Puede decir bastante sobre los datos



Distribuciones de probabilidad

- Función que entrega las probabilidades que puede tomar cada valor de la variable aleatoria X

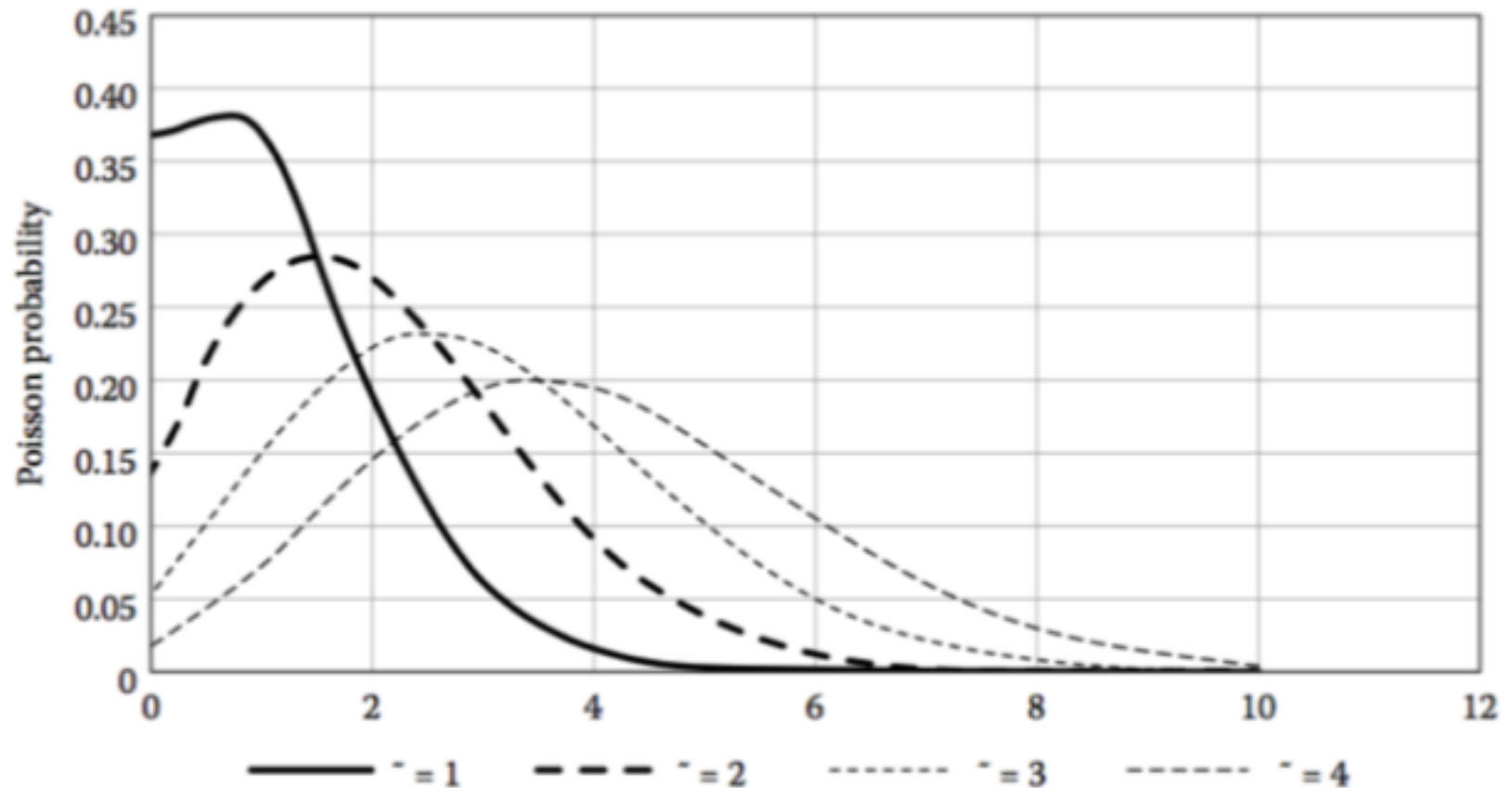
Distribución de Poisson

- Distribución de probabilidad discreta que se aplica a las ocurrencias de algún suceso durante un intervalo específico

$$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$

- X es el número de veces que el suceso ocurre en el intervalo
- μ es el promedio de veces que el suceso ocurre en el intervalo

Distribución de Poisson



Ejemplo

El número de correos que llegan a un servidor en un período de 15 minutos puede ser descrito como una distribución Poisson con una media de 2.

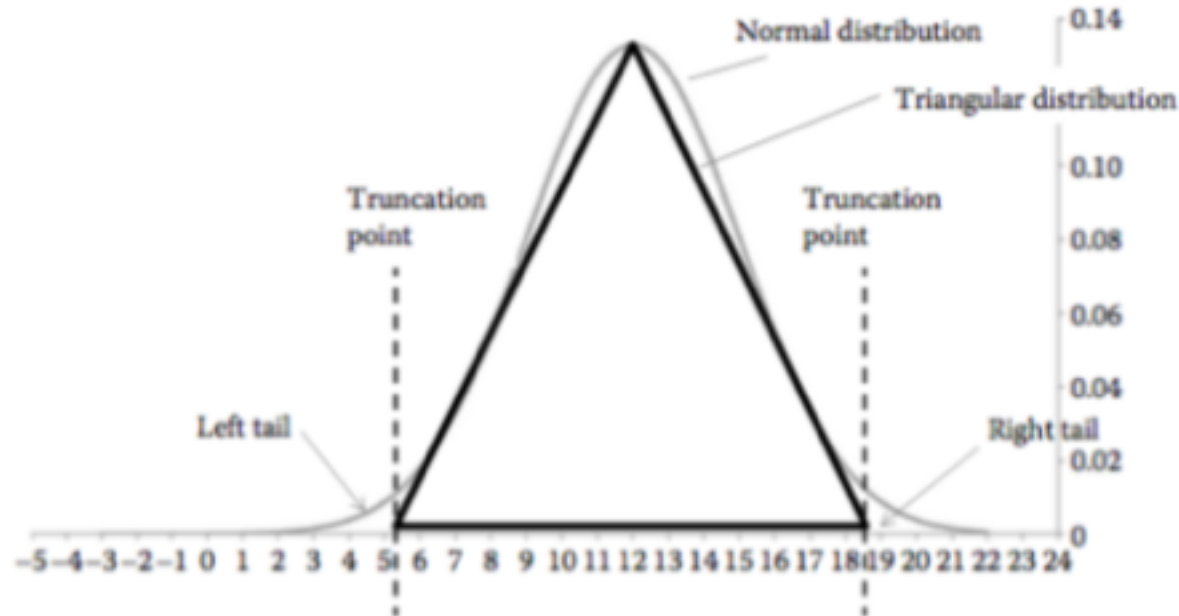
- a) ¿Cuál es la probabilidad de que no se reciba ningún mensaje en un intervalo de 15 minutos?
- b) ¿Cómo es la distribución de mensajes que llegan en una hora?

a) $p(x=0) = e^{-2} = 0.135$

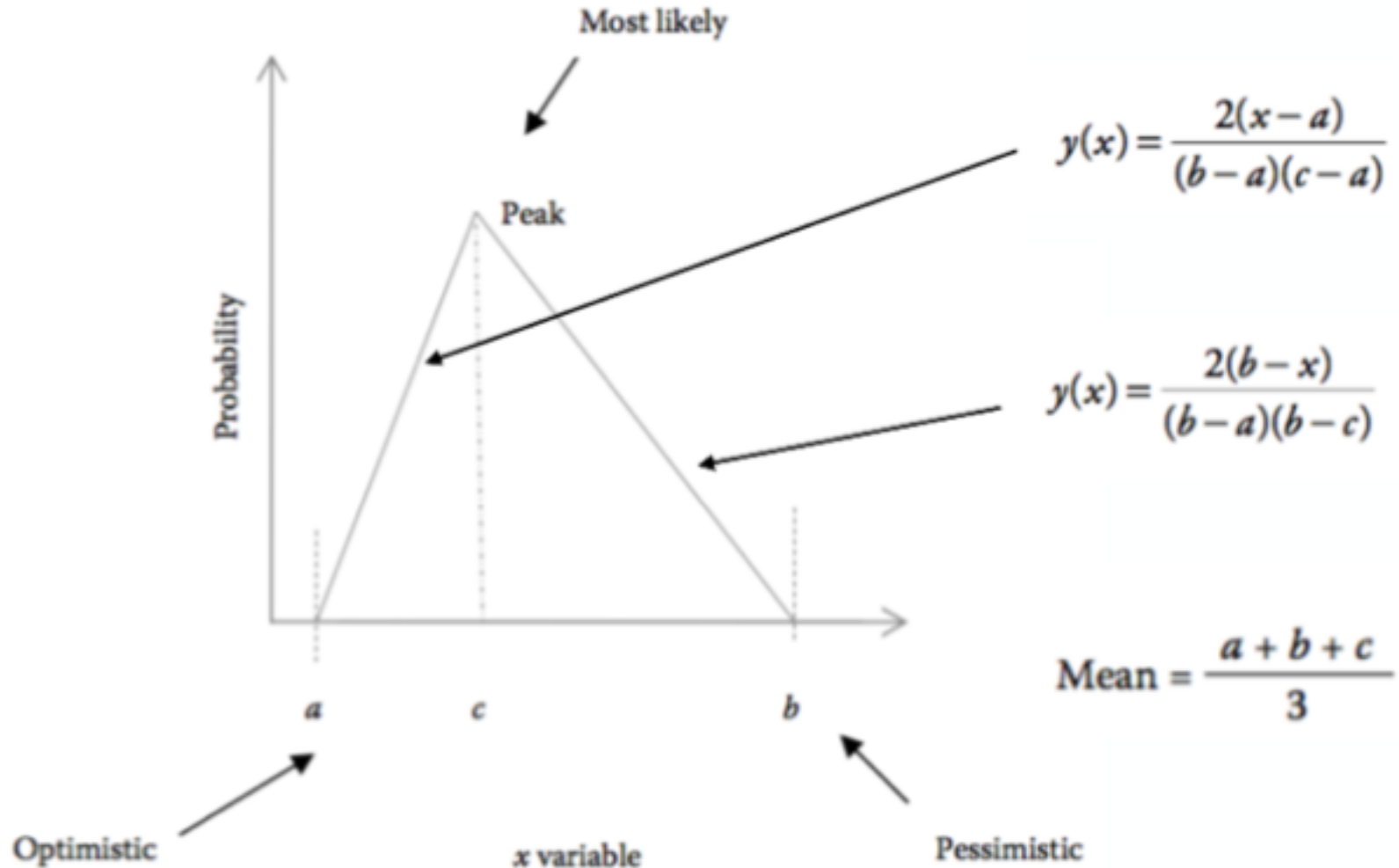
b) Distribución Poisson con media 8

Distribución triangular

- Aproximación simple de una distribución normal



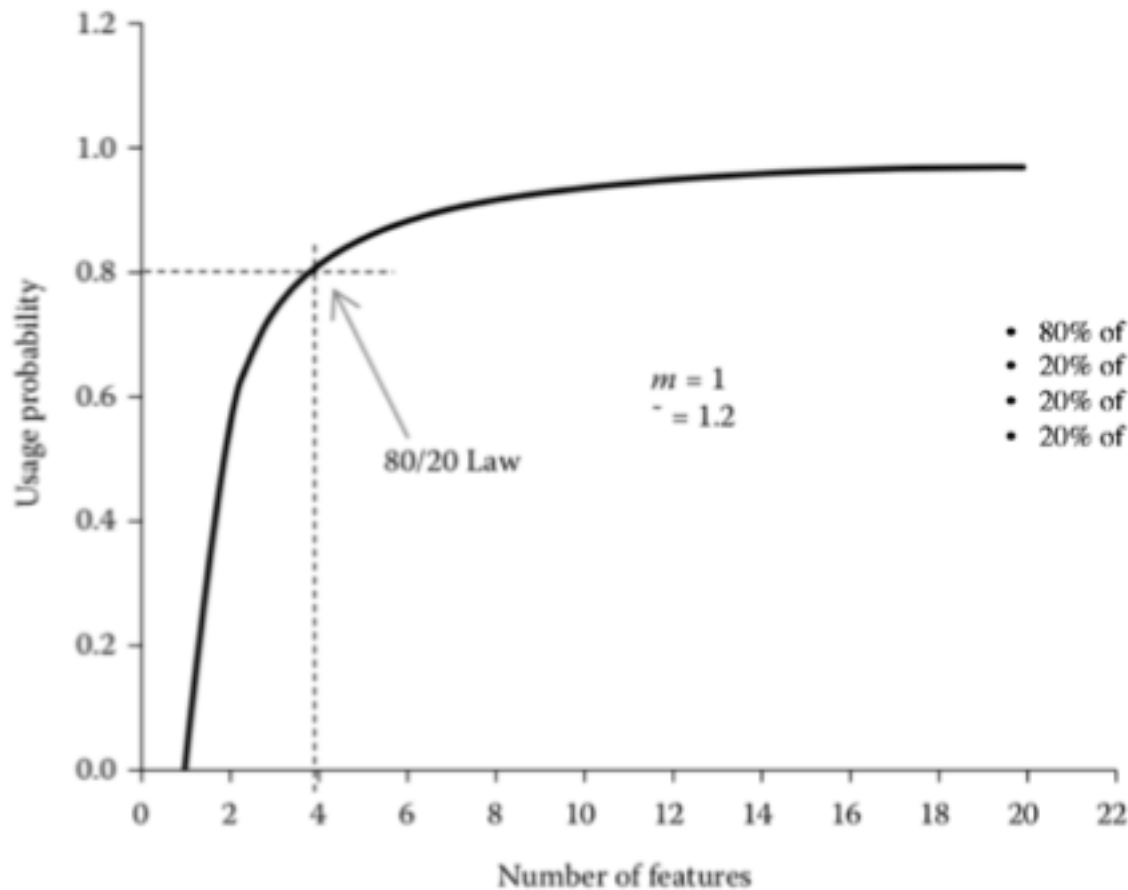
Distribución triangular



Ejemplo

- Estimación duración de un proyecto
 - Optimista: 25 días
 - Pesimista: 50 días
 - Más probable: 30 días
- Media: $(25 + 50 + 30)/3 = 35$ días
- PERT: $(25 + 4 * 30 + 50)/6 = 32.5$ días

Ley de Pareto



- 80% of errors and crashes come from 20% of bugs
- 20% of software components contain 80% of defects
- 20% of defects cause 80% of down time
- 20% of test cases capture 80% of defects

SQA Estadístico

- Recolectar datos sobre errores y defectos
- Trazar para cada error su causa
- Seleccionar el 20% de causas que generan el 80% de los errores (ley de Pareto)
- Enfocarse en evitar/corregir esas causas

Ejemplo

Error	Total		Serious		Moderate		Minor	
	No.	%	No.	%	No.	%	No.	%
IES	205	22%	34	27%	68	18%	103	24%
MCC	156	17%	12	9%	68	18%	76	17%
IDS	48	5%	1	1%	24	6%	23	5%
VPS	25	3%	0	0%	15	4%	10	2%
EDR	130	14%	26	20%	68	18%	36	8%
ICI	58	6%	9	7%	18	5%	31	7%
EDL	45	5%	14	11%	12	3%	19	4%
IET	95	10%	12	9%	35	9%	48	11%
IID	36	4%	2	2%	20	5%	14	3%
PLT	60	6%	15	12%	19	5%	26	6%
HCI	28	3%	3	2%	17	4%	8	2%
<u>MIS</u>	<u>56</u>	<u>6%</u>	<u>0</u>	<u>0%</u>	<u>15</u>	<u>4%</u>	<u>41</u>	<u>9%</u>
Totals	942	100%	128	100%	379	100%	435	100%

- IES - especificación errónea o incompleta
- MCC - malentendido en requerimiento del cliente
- EDR - error en representación de datos
- IID - documentación incompleta
- ICI - interfaz inconsistente
- IID - error en transf de diseño a código
- EDL - error en diseño lógico
- HCI - interfaz humano computador

Teorema de Bayes

- La probabilidad de un evento futuro está influenciada por la historia

$$p(A|B) = \frac{p(A \text{ and } B)}{P(B)}$$

- Estimaciones de costos o productividad
 - juicio a priori de experto se combina con información histórica



Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación

Clase 24

Métodos estadísticos en ingeniería de software

IIC2143 - Ingeniería de Software
Sección 1

Rodrigo Saffie

rasaffie@uc.cl

20 de junio de 2018