

---

# 基本情報技術 I

## 4. 確率と統計 菊池浩明

# 講義目標

---

## ■ 教科書

### □ 1-3 応用数学

- » 1. 確率
- » 2. 確率分布
- » 3. 回帰直線

# 宿題(次回小テスト範囲)

---

- 1章練習問題 (p.57)

□11, 12

# 1. 確率

---

## ■ 組み合わせ

□  $n$ 個の要素から $r$ 個の要素を取り出す場合の数.

$${}_nC_r = {}_nP_r / r! = \binom{n}{r} = \frac{n!}{(n-r)!r!}$$

□ 例) 8人から3人を選ぶ組み合わせの数

$$\binom{8}{3} =$$

# 積事象

---

## ■ 独立事象

- 事象Aと事象Bが独立に生じる.
- 例) 2個のさいころ共に偶数となる確率.  
 $P(A \cap B) = P(A)P(B) = 1/4$   
A=1個目が偶数, B=2個目が偶数

## ■ 事象

- Bの生起がAの結果に影響を受ける.
- 例) 白3個, 赤5個の壺から2個の玉を取る.  
 $P(A \cap B) = P(2個とも赤) = P(A) P(B|A) =$   
A = 1個目が赤,  $P(A) = 5/8$   
B = 2個目が赤,  $P(B|A) = A$ による 確率 =  
 $P(B) = P(B|A)P(A) + P(B|\sim A)P(\sim A) = 4/7 * 5/8 + 5/7 * 3/8 =$

# 条件付き確率と同時確率

## ■ 同時確率

	A	~A	計
B	$\frac{20}{56}$ $P(\quad)$	$\frac{15}{56}$	$\frac{5}{8}$ $P(B)$
~B	$\frac{15}{56}$	$\frac{6}{56}$	$\frac{3}{8}$
計	$\frac{5}{8}$ $P(\quad)$	$\frac{3}{8}$	1.0

## ■ 条件付き確率

	A	~A
B	$\frac{4}{7}$ $P(\quad)$	$\frac{5}{7}$ $P(B \sim A)$
~B	$\frac{3}{7}$ $P(\sim B A)$	$\frac{2}{7}$
計		1.0

# 和事象, 余事象

---

## ■ 和事象

□  $P(A \cup B) = P(A) + P(B) - P(\quad)$

□ 例) 2個のどちらかが赤となる確率

$$P(A \cup B) = 5/8 + 5/8 - 5/14 =$$

## ■ \_\_ 事象

□ 事象Aが起きない事象

□ 例)  $C =$  2個とも白になる事象,  $P(C) =$

□  $\bar{C} = A \cup B =$  少なくともどちらかが赤となる事象,  
 $P(\bar{C}) =$

# ベイズの定理\*

---

## ■ 例(ほとんど出題されない?)

□ABのサイコロでE=1の目が出る事象の確率が $P(E|A) = 3/10$ ,  $P(E|B)=3/5$ である時, Eの時Aであった確率 $P(A|E)$ を求めよ.

□ベイズの定理

$$\begin{aligned} P(A|E) &= P(A \cap E) / P(E) = P(\quad) P(A) / P(E) \\ &= 3/10 * 1/2 / (9/20) = 1/3 \end{aligned}$$

$$P(A)=P(B)=1/2$$

$$\begin{aligned} P(E) &= P(E|A)P(A) + P(E|B)P(B) \\ &= (3/10+3/5)/2 = 9/20 \end{aligned}$$



# 確率に関する問題例

---

- (1) 赤, 白, 黄が3個ずつ入った壺がある.  
3つ出したらすべて白の確率.
- (2) AとBの文字を $n$ 個ならべて, 60個以上の数を表すには $n$ は何個必要か?
- (3) 0と1を組み合わせてできる長さ1以上7文字以下の文字列の総数.

## 2. 確率分布

### ■ 統計値

□ 平均  $\bar{x} = \mu = 1/n \sum x_i$

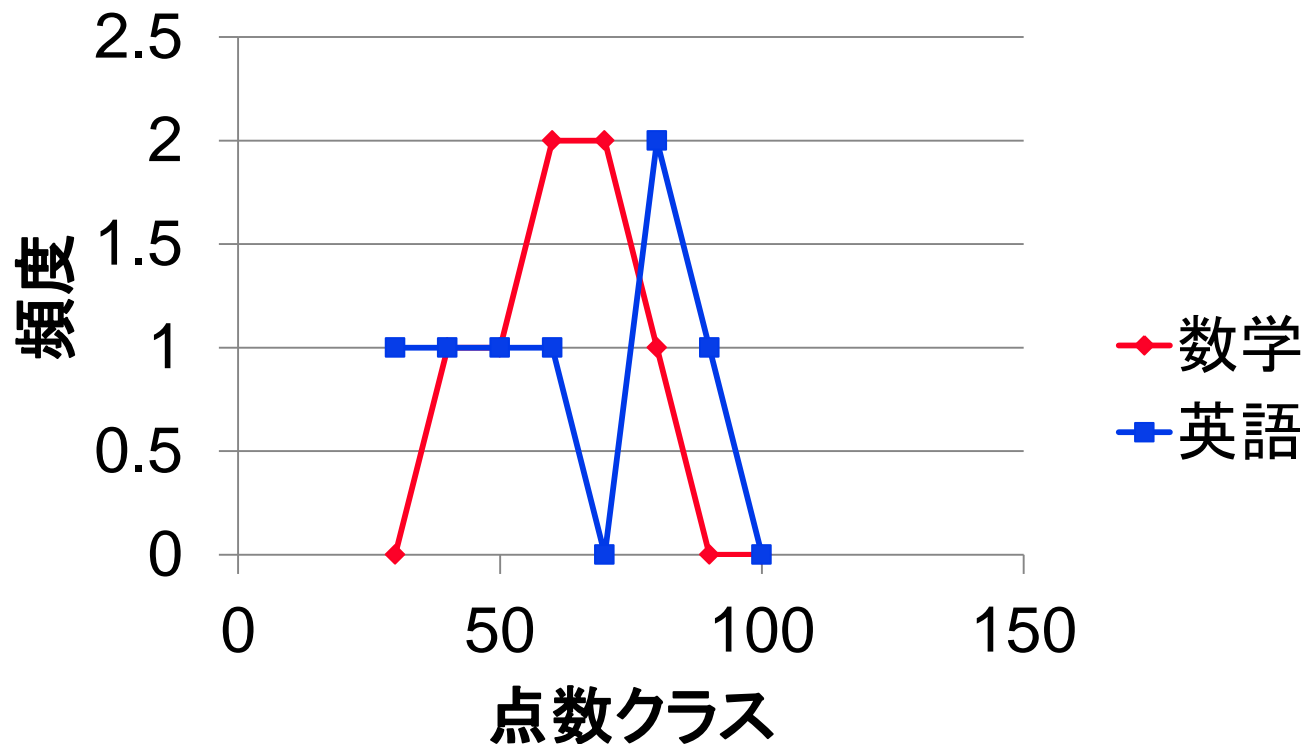
□ 分散  $\sigma^2 = 1/n \sum (x_i - \mu)^2$

□ 標準偏差  $\sigma = \sqrt{(\sigma^2)}$

	60	60
	55	45
	65	75
	70	80
	50	40
	80	90
	40	30
平均	60.00	60.00
	150	435.7
標準偏差		20.9

## ■ 度数分布表

□ 確率分布: 確率=頻度/N



# 正規分布

## ■ 定義

□ 自然現象によく観られる確率分布 . \_\_\_\_関数

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

□ 標準正規分布

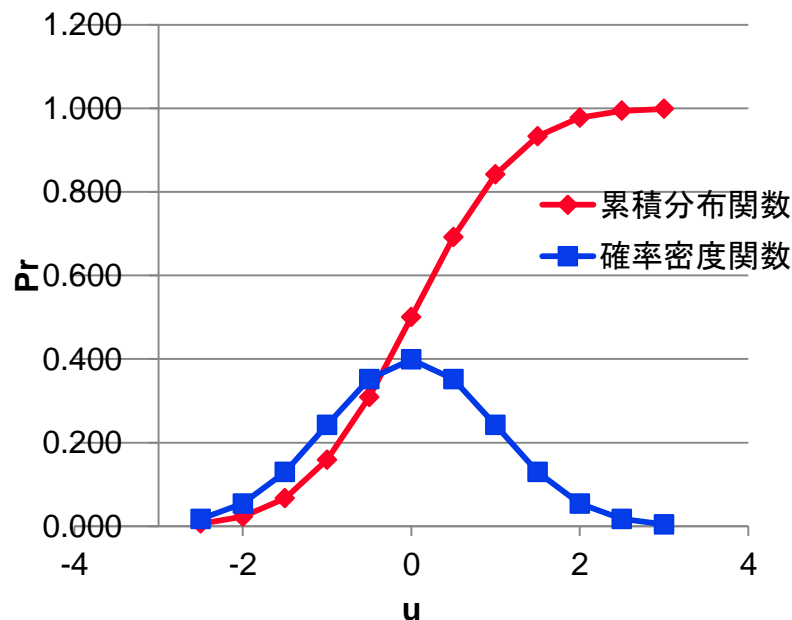
$$\mu = , \sigma =$$

$$N(0, 1)$$

□ 累積確率 \_\_\_\_関数

$$P(X < t)$$

$$= \int_A f_X(x) dx$$



# 例題

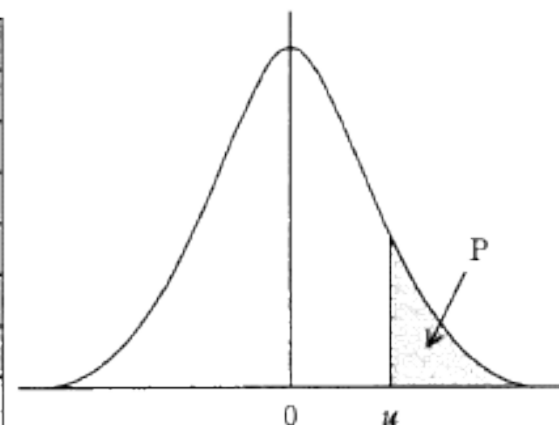
- 平均5.2kg, 標準偏差0.1kgで分布している製品がある. 5.0kg未満が不合格品とする. その割合はいくらか？

□  $u = (x - \mu) / \sigma = (5.0 - 5.2) / 0.1 = -2$  より,  $P(2) = 0.023 = 1 - \Pr(U < 2)$ . 答            %

- $X \geq 5.3\text{kg}$ となる製品の割合は？

標準正規分布表

$u$	P
0.0	0.500
0.5	0.309
1.0	0.159
1.5	0.067
2.0	0.023
2.5	0.006
3.0	0.001



# 参考) 偏差値\*

---

## ■ 定義

$$T_i = \frac{10(x_i - \mu_x)}{\sigma_x} + 50$$

□ 例) 算数  $\mu=60$ ,  $\sigma=12.2$ ,  $x=40$ 点の時 ,  
 $T=10(40 - 60)/12.2 = \underline{33.6}$

□ 英語  $T = 10(40-60)/20.9 = \underline{40.4}$

□ (            の方がヤバい)

- FMSの偏差値が60点である. 合格するのは100人中何位までか？

# 3. 統計

## ■ \_\_\_\_\_ 係数 (ピアソンの積率相関係数)

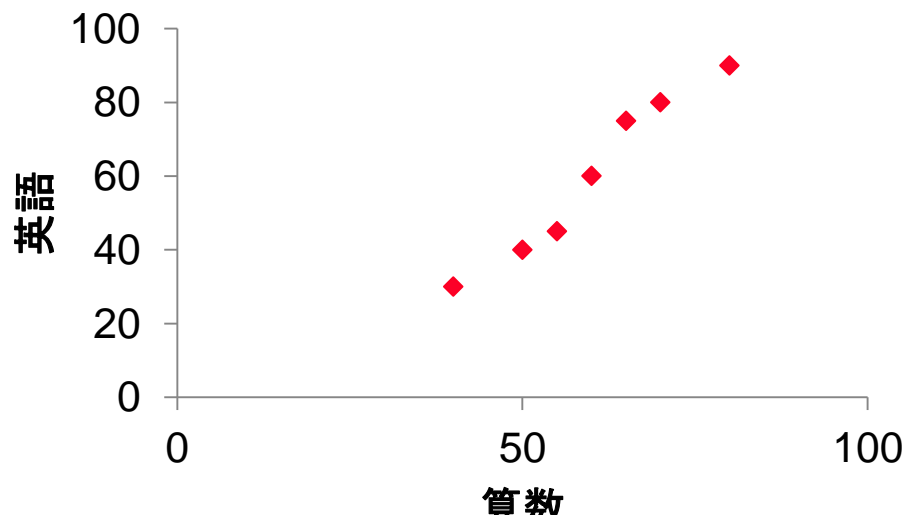
□  $(x_i, y_i)$  の組が与えられてるとき

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \sum (x_i - \mu_x)(y_i - \mu_y) / \sigma_x \sigma_y$$

$y_i = x_i$  の時,  $r = \sum (x_i - \mu_x)^2 / \sigma_x^2 = \underline{\hspace{1cm}}$  (最大値)

例) 算数と英語の  
点数散布図の時,

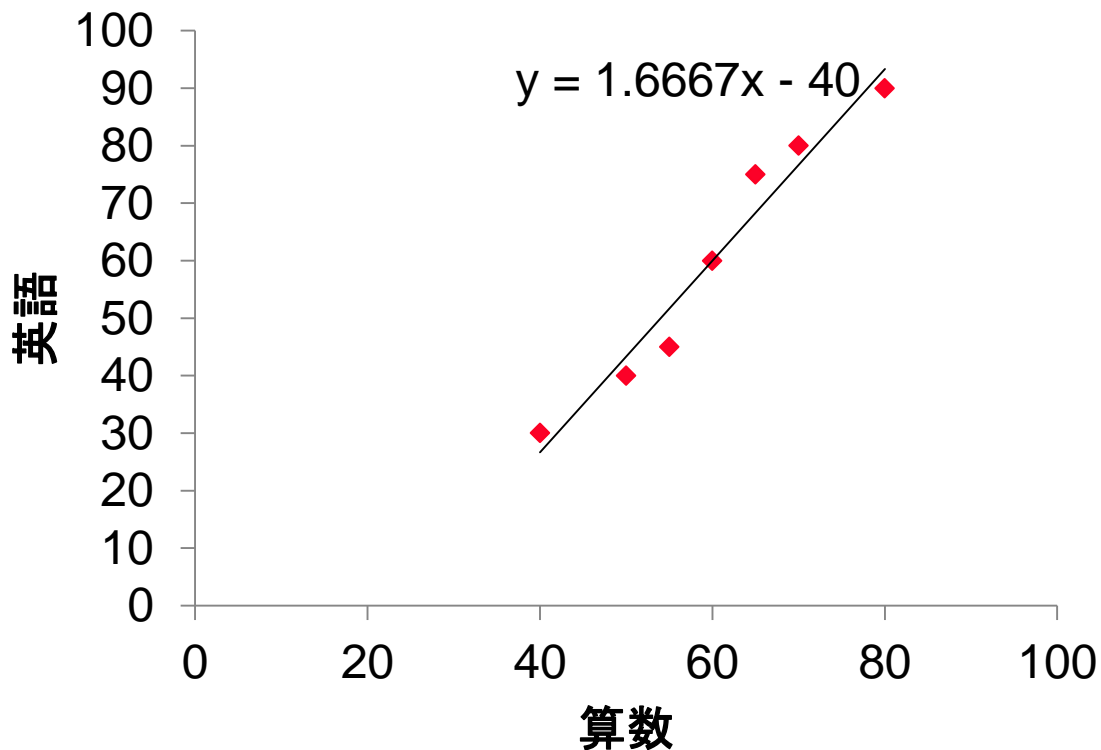
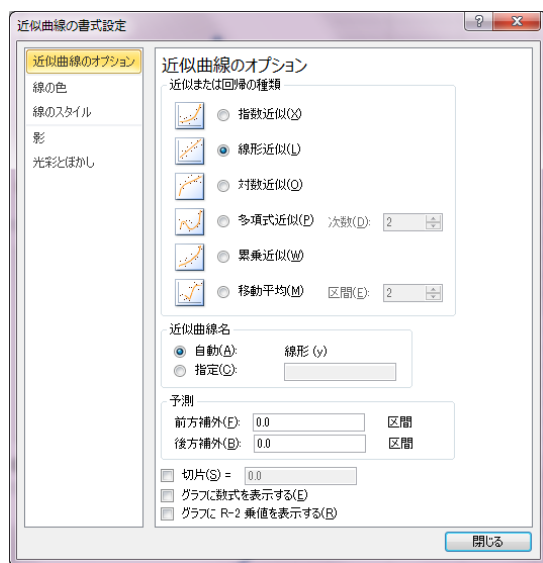
$$R = \boxed{0.977897823}$$



# 回帰直線

## ■ 最小二乗法

### □ 「近似曲線の追加」





# 参考) 最小二乗法\*

---

- アルゴリズム

- 入力:  $(x_1, y_1), \dots, (x_n, y_n)$

- あてはめ多項式:  $f(x) = a + bx$

- 誤差の総和  $S = \sum (f(x_i) - y_i)^2$

- $dS/da = 0$

- $dS/db = 0$  の連立方程式を解く.

# まとめ

---

- 二つのサイコロの様に独立して生じる事象を(            ), そうでないものを(            )という.  
 $P(A|B)$ を(    )による(            )確率という.
- 標準正規確率分布とは, 平均が(            )で(            )が1.0の正規分布である. 標準偏差を二乗すると(            )になる.
- 相関係数は最大値(            ), 最小値(            )で二つの分布の相関を表す. 別名(            )の積率相関係数という.