

# Biometryczna Identyfikacja Tożsamości

## Wykład 10: Statystyczna ocena systemów biometrycznych

Adam Czajka

Wykład na Wydziale Elektroniki i Technik Informatycznych  
Politechniki Warszawskiej

Semestr letni 2015

## Wykład 10: Statystyczna ocena systemów biometrycznych

### Problem ewaluacji w biometrii

Modelowanie niepewności

Ocena procesów pozyskiwania próbek i rejestracji

Ocena zgodności próbek

Ocena systemów biometrycznych

Ocena rozrzutu parametrów

Wybrane uwagi praktyczne

## Problem ewaluacji systemów biometrycznych

Zaprojektowaliśmy i wykonaliśmy system biometryczny (urządzenia, algorytmy, procesy decyzyjne, itd.). Jesteśmy z niego dumni :-)

Znaleźliśmy odbiorcę, i niedługo nadejdzie czas przekazania systemu. Jakie pytania będą (powinny) nas nurtować?

## Problem ewaluacji systemów biometrycznych

1. Pomiary własności istot żywych → niepewność:  
jak uwzględnić niepewność w opisie działania?
2. Ocena jakości systemu bazuje na eksperymencie:  
jaki eksperyment należy wykonać?
3. Decyzje zależą od progów decyzyjnych (odrzuć próbki złej jakości, odrzuć deklarowane tożsamości):  
jakie powinny być wartości tych progów?
4. Eksperymenty wymagają biometrycznych baz danych:  
jaka powinna być liczność próbek w takiej bazie?
5. Wyliczyłem podstawowe parametry systemu (np. EER):  
jakie poziomy błędów mogę gwarantować mojemu odbiorcy?

## Ocena biometrii wg ISO/IEC 19795

1. Ocena technologiczna: ocena off-line jednego lub więcej algorytmu dla tej samej modalności biometrycznej przy zastosowaniu istniejącej lub specjalnie utworzonej próby
2. Ocena w scenariuszu zastosowań: ocena off-line lub on-line w której funkcjonowanie systemu biometrycznego jest oceniane dla określonego scenariusza zastosowań z określoną grupą docelową
3. Ocena w warunkach operacyjnych: ocena on-line funkcjonowania systemu w trakcie normalnej eksploatacji w warunkach, do których był zaprojektowany

# Ocena biometrii wg ISO/IEC 19795

## Ocena technologiczna

1. **Ocena technologiczna**: ocena **off-line** jednego lub więcej algorytmu dla tej samej modalności biometrycznej przy zastosowaniu **istniejącej lub specjalnie utworzonej próby**
  - najczęściej ocena wybranego komponentu systemu (np. algorytmu lub urządzenia)
  - próba ustalona, pobrana przez “uniwersalny” sensor
  - wynik zależy od środowiska i populacji, z której pobrano próbę
  - próba nie powinna być znana twórcom testowanego komponentu
  - wyniki powtarzalne
  - znaczne niedoszacowanie błędów w stosunku do kompletnego systemu wykorzystującego testowany komponent i pracującego w warunkach operacyjnych

**Przykład**: ocena weryfikacji algorytmu rozpoznawania tęczówki firmy XYZ

# Ocena biometrii wg ISO/IEC 19795

## Ocena w scenariuszu zastosowań

2. Ocena w scenariuszu zastosowań: ocena off-line lub on-line w której funkcjonowanie systemu biometrycznego jest oceniane dla określonego scenariusza zastosowań z określoną grupą docelową
  - testowanie kompletnego systemu
  - środowisko testowe bliskie planowanemu
  - każdy system powinien mieć własny sensor
  - konieczne to samo środowisko i populacja do porównań systemów
  - powtarzalność pod warunkiem kontroli scenariusza
  - zwykle niedoszacowanie błędów w porównaniu do warunków operacyjnych

**Przykład:** ocena weryfikacji systemu rozpoznawania tęczy na symulowanym przejściu granicznym

# Ocena biometrii wg ISO/IEC 19795

## Ocena w warunkach operacyjnych

3. Ocena w warunkach operacyjnych: ocena on-line funkcjonowania systemu w trakcie normalnej eksploatacji w warunkach, do których był zaprojektowany
  - autentyczna obsługa systemu, warunki, użytkownicy, itp.
  - warunki mogą być nienadzorowane
  - warunki mogą się zmieniać w nieprzewidziany sposób (zalecana rejestracja zmian)
  - brak powtarzalności wyników
  - najlepsza ocena funkcjonowania systemu (choć nie zawsze możliwa i często kosztowna)

**Przykład:** program pilotażowy weryfikacji tęczyówki na przejściu granicznym dla osób często podróżujących



## Wykład 10: Statystyczna ocena systemów biometrycznych

Problem ewaluacji w biometrii

### **Modelowanie niepewności**

Ocena procesów pozyskiwania próbek i rejestracji

Ocena zgodności próbek

Ocena systemów biometrycznych

Ocena rozrzutu parametrów

Wybrane uwagi praktyczne

# Próba losowa

## 1. Populacja generalna $X$

- zmienna losowa (jedno- lub wielowymiarowa), której realizacje dostarczają nam wszystkich interesujących nas informacji o badanym zjawisku

## 2. Próba losowa $X_1, \dots, X_n$ zmiennej losowej $X$

- wynik wielokrotnej obserwacji (próbki) zm. los.  $X$
- jeśli próbki  $X_1, \dots, X_n$  są niezależne (kolejne obserwacje nie są od siebie w jakikolwiek sposób uzależnione) mówimy, że próba losowa jest **prosta** (ang. *i.i.d. = independent and identically-distributed*)

### W biometrii:

- **populacja generalna** = dane wszystkich potencjalnych użytkowników badanego systemu biometrycznego
- **próba losowa** = podzbiór danych użytkowników systemu (wykorzystywany do oceny danego systemu)
- **próbka** (wynik obserwacji zmiennej losowej) = próbka biometryczna

## Tworzenie próby losowej w biometrii

1. Prezentacja (ang. *presentation*) charakterystyki biometrycznej
2. Podejście (ang. *attempt*) złożone z jednej lub kilku prezentacji (np. w celu doboru najlepszej)
3. Transakcja (ang. *transaction*) złożona z jednego lub kilku podejść (zależy od polityki systemu)

# Probabilistyka i statystyka

## 1. Probabilistyka

- rozkłady prawdopodobieństwa zmiennych losowych są znane
- **obliczanie** prawdopodobieństw zdarzeń na bazie znanych rozkładów

## 2. Statystyka

- dowolna funkcja próby losowej  $g(X_1, \dots, X_n)$
- statystyka jest zmienną losową
- brak założeń o pełnej znajomości rozkładów prawdopodobieństwa zmiennych losowych
- **wyciąganie wniosków** dotyczących badanej cechy w populacji generalnej na podstawie próby losowej

# Estymacja

1. Szacowanie nieznanego parametru rozkładu za pomocą wartości statystyki (konkretnej liczby)
2. Estymator: dowolna statystyka  $\hat{\theta}_n = g(X_1, \dots, X_n)$ , której wartości wykorzystujemy do oceny nieznanego parametru  $\theta$
3. Pożądane cechy estymatorów
  - nieobciążoność:  $\mathcal{E}\hat{\theta} = \theta$
  - zgodność:  $\bigwedge_{\epsilon > 0} \lim_{n \rightarrow \infty} \mathcal{P}(|\hat{\theta}_n - \theta| < \epsilon) = 1$

## Przykłady estymatorów w biometrii

- średnia arytmetyczna  $\hat{m}$  z wyników porównań tych samych tęczówek (estymator wartości oczekiwanej  $m$  zmiennej losowej  $X$ )
- FNMR (częstość fałszywego niedopasowania)
- FMR (częstość fałszywego dopasowania)

## Wykład 10: Statystyczna ocena systemów biometrycznych

Problem ewaluacji w biometrii

Modelowanie niepewności

**Ocena procesów pozyskiwania próbek i rejestracji**

Ocena zgodności próbek

Ocena systemów biometrycznych

Ocena rozrzutu parametrów

Wybrane uwagi praktyczne

# Niepowodzenie pozyskania próbki

## ang. Failure to Acquire (FTA)

1. Powody niemożności pozyskania próbki
  - dane biometryczne nie mogły być zaprezentowane (choroba, zranienie, prezentacja niezgodna z regułami)
  - zawiódł algorytm lub urządzenie (przekroczenie czasu, błąd)
  - zawiódło przetwarzanie próbki (np. segmentacja, usuwanie zakłóceń, wyznaczanie cech)
  - wyliczone cechy nie miały wystarczającej jakości
2. Empiryczny estymator prawdopodobieństwa niemożności pozyskania próbki: **częstość FTA**

$$\text{FTA} = \frac{\text{liczba podejść zakończonych niepowodzeniem}}{\text{liczba wszystkich podejść}}$$

# Niepowodzenie rejestracji

ang. **Failure to Enroll (FTE)**

1. Powody niemożności rejestracji
  - jak dla FTA, oraz dodatkowo:
  - próbka nie została zaakceptowana w wyniku próbnej weryfikacji
2. Empiryczny estymator prawdopodobieństwa niemożności rejestracji: **częstość FTE**

$$\text{FTE} = \frac{\text{liczba rejestracji zakończonych niepowodzeniem}}{\text{liczba wszystkich rejestracji}}$$

3. Surowsze wymagania na etapie rejestracji **zwiększają jakość rozpoznawania** (w szczególności mniejsze FMR, FNMR), ale **zmniejszają liczbę zarejestrowanych wzorców** (większe FTE)



## Wykład 10: Statystyczna ocena systemów biometrycznych

Problem ewaluacji w biometrii

Modelowanie niepewności

Ocena procesów pozyskiwania próbek i rejestracji

**Ocena zgodności próbek**

Ocena systemów biometrycznych

Ocena rozrzutu parametrów

Wybrane uwagi praktyczne

# Rodzaje i dopasowanie próbek

## 1. Rodzaje próbek biometrycznych

- **próbki własne**: pochodzące z tej samej klasy
- **próbki obce** lub **fałszerstwa bezwysiłkowe**: pochodzące z innej klasy w takiej formie, w jakiej zostały pozyskane
- **fałszerstwa zaawansowane**: spreparowane specjalnie, aby imitować próbkę z badanej klasy

# Rodzaje i dopasowanie próbek

## 2. Dopasowanie próbek

- stopień podobieństwa/niepodobieństwa lub dopasowania/niedopasowania
  - podobieństwo = 0 → próbki całkowicie niedopasowane
  - niepodobieństwo = 0 → próbki identyczne
- nie zawsze jest łatwe przejście między podobieństwem i niepodobieństwem

## 3. Decyzja

- dopasowanie/niedopasowanie na podstawie porównania stopnia dopasowania/niedopasowania z progiem decyzyjnym

# Testowanie hipotez

## 1. Hipotezy

- hipoteza zerowa ( $H_0$ )
  - $H_0$  weryfikujemy
  - za  $H_0$  przyjmujemy tę hipotezę, dla której chcemy szukać podstaw do jej odrzucenia
  - przyjęcie  $H_0$  nie oznacza jej prawdziwości, wynika jedynie z braku podstaw do jej odrzucenia
- hipoteza alternatywna ( $H_1$ )
  - $H_1$  jesteśmy skłonni przyjąć w momencie odrzucenia  $H_0$

## 2. Rozpatrujemy najprostszyp przypadk

- $H_0$  i  $H_1$  są hipotezami prostymi, czyli:
  - precyzują rozkłady z których losowane są próby
  - możemy zakładać znajomość tych rozkładów

# Testowanie hipotez

## Przykłady hipotez

- $H_0$  (zerowa) prosta
  - tęczówka X jest tęczówką A. Czajki
- $H_1$  (alternatywna) prosta
  - tęczówka X jest tęczówką J. Kowalskiego
  - tęczówka X jest tęczówką osoby z ustalonej grupy N osób
- $H_1$  (alternatywna) złożona
  - tęczówka X nie jest tęczówką A. Czajki

## Testowanie hipotez

### 3. Test statystyczny

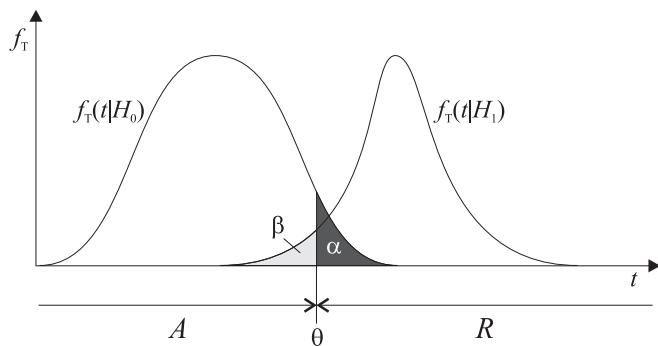
- metoda przyporządkowująca realizacjom próby losowej decyzje przyjęcia bądź odrzucenia weryfikowanej hipotezy (z ustalonym prawdopodobieństwem)

### 4. Statystyka testowa $T$

- statystyka wykorzystywana do oceny hipotezy  $H_0$
- zbiór wartości statystyki  $T$  dzielimy na dwa dopełniające się podzbiory
  - $R$ : zbiór krytyczny hipotezy  $H_0$  ( $H_0$  odrzucamy jeśli  $T \in R$ )
  - $A$ : zbiór przyjęć hipotezy  $H_0$  ( $H_0$  przyjmujemy jeśli  $T \in A$ )

Przykład statystyki testowej: znormalizowana odległość Hamminga binarnych kodów tęczówki

## Testowanie hipotez



$H_0$ : realizacja pochodzi z rozkładu "0"  $\rightarrow$  statystyka testowa  $T$  ma rozkład  $f_T(t|H_0)$

$H_1$ : realizacja pochodzi z rozkładu "1"  $\rightarrow$  statystyka testowa  $T$  ma rozkład  $f_T(t|H_1)$

# Testowanie hipotez

Podejmując decyzje zawsze popełniamy błąd

		decyzja	
		brak podstaw do odrzucenia	odrzucenie
hipoteza	prawdziwa	OK ( $1 - \alpha$ )	błąd I rodzaju ( $\alpha$ )
	falszywa	błąd II rodzaju ( $\beta$ )	OK ( $1 - \beta$ )

gdzie  $1 - \alpha$ : poziom istotności,  $1 - \beta$ : moc testu



# Testowanie hipotez

## Związek z decyzjami biometrycznymi

		decyzja	
		zgodność z wzorcem	niezgodność z wzorcem
próbka	zgodna z wzorcem	OK (1-FNMR)	FNM (FNMR)
	niezgodna z wzorcem	FM (FMR)	OK (1-FMR)

gdzie FNM oznacza błędne niedopasowanie, FM oznacza błędne dopasowanie, natomiast FNMR oraz FMR są estymatorami punktowymi prawdopodobieństwa popełnienia błędu, odpowiednio, FNM i FM

## Ocena fałszywego niedopasowania

ang. False Non-Match (FNM), False Non-Match Rate (FNMR)

1. FNM: błędne uznanie próbki za niezgodną z odpowiednim wzorcem klasy przy prezentacji próbek zgodnych z wzorcem
2. FNM odpowiada błędowi I rodzaju w testowaniu hipotez statystycznych (odrzućenie hipotezy prawdziwej)
3. Empiryczny estymator prawdopodobieństwa błędnego niedopasowania próbki: częstość błędnych niedopasowań, obliczany tylko dla próbek zgodnych z wzorcami

$$\text{FNMR} = \frac{\text{liczba podejść uznanych za niezgodne z wzorcami}}{\text{liczba wszystkich podejść}}$$

4. FNMR jest funkcją progu decyzyjnego

# Ocena fałszywego dopasowania

ang. False Match (FM), False Match Rate (FMR)

1. FM: błędne uznanie próbki za **zgodną z odpowiednim wzorcem** klasy przy prezentacji próbek **niezgodnych z wzorcem**
2. FM odpowiada **błędowi II rodzaju** w testowaniu hipotez statystycznych (akceptacja hipotezy fałszywej)
3. Empiryczny estymator prawdopodobieństwa błędnego dopasowania próbki: **częstość błędnych dopasowań**, obliczany **tylko dla próbek obcych i fałszerstw bezwysiłkowych**

$$\text{FMR} = \frac{\text{liczba podejść uznanych za zgodne z wzorcami}}{\text{liczba wszystkich podejść}}$$

4. Do wyliczania FMR **nie włącza się fałszerstw zaawansowanych**
5. FMR, podobnie jak FNMR, jest **funkcją progu decyzyjnego**

## Wykład 10: Statystyczna ocena systemów biometrycznych

Problem ewaluacji w biometrii

Modelowanie niepewności

Ocena procesów pozyskiwania próbek i rejestracji

Ocena zgodności próbek

**Ocena systemów biometrycznych**

Ocena rozrzutu parametrów

Wybrane uwagi praktyczne

## Różnice w ocenie podejścia i transakcji

1. Ocena systemu polega na badaniu **transakcji a nie pojedynczych podejść**
2. Ocena dotyczy **całego systemu**, zatem powinna uwzględniać:
  - możliwość niepozyskania próbki FTA (wtedy próba weryfikacja kończy się odrzuceniem)
  - politykę decyzyjną systemu, w tym liczbę dopuszczalnych podejść dla każdej transakcji
3. Transakcję będziemy nazywać **autentyczną**, jeśli prezentowana próbka pochodzi z rozpatrywanej klasy; w przeciwnym razie nazwiemy ją **fałszywą**

## Ocena systemów weryfikacji

### Częstość fałszywego odrzucenia (ang. False Rejection Rate (FRR))

1. FR: fałszywe odrzucenie (ang. false rejection); **odrzucenie autentycznej transakcji**, tzn. negatywna weryfikacja próbki pochodzącej z właściwej klasy
2. Empiryczny estymator prawdopodobieństwa błędnego odrzucenia próbki to **częstość błędnych odrzuceń**:

$$FRR = \frac{\text{liczba odrzuconych autentycznych transakcji}}{\text{liczba wszystkich transakcji}}$$

3. FRR jest funkcją **polityki decyzyjnej**; np. dla systemu z jednym podejściem w transakcji:  $FRR = FTA + (1-FTA) FNMR$

# Ocena systemów weryfikacji

## Częstość fałszywej akceptacji (ang. False Acceptance Rate (FAR))

1. FA: fałszywa akceptacja (ang. false acceptance); **akceptacja fałszywej transakcji**, tzn. pozytywna weryfikacja próbki pochodzącej z innej klasy
2. Empiryczny estymator prawdopodobieństwa błędnej akceptacji próbki to **częstość błędnych akceptacji**:

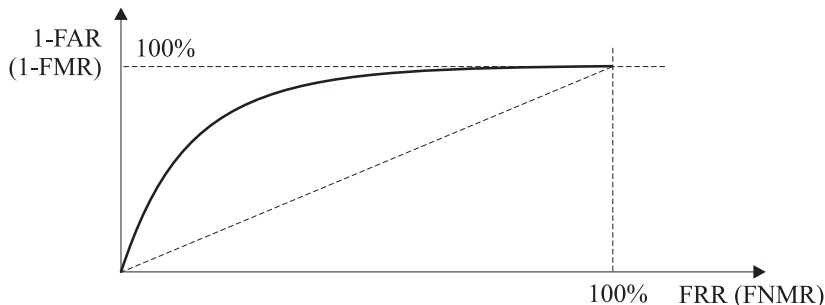
$$\text{FAR} = \frac{\text{liczba zaakceptowanych fałszywych transakcji}}{\text{liczba wszystkich transakcji}}$$

3. FAR, podobnie jak FRR, jest funkcją **polityki decyzyjnej**; np. dla systemu z jednym podejściem w transakcji:  
 $\text{FAR} = (1 - \text{FTA}) \text{FMR}$
4. W obliczaniu FAR **nie uwzględnia się niepozyskanych próbek**

# Ocena systemów weryfikacji

## Krzywa ROC (ang. Receiver Operating Characteristic)

1. Pierwotnie wykorzystywana do analizy sygnałów radarowych podczas II wojny światowej; obecnie często wykorzystywana w analizie testów medycznych
2. Połączenie estymatorów FAR i FRR (lub FMR i FNMR) w jedną **krzywą parametryczną** (parametrem jest wartość progu)

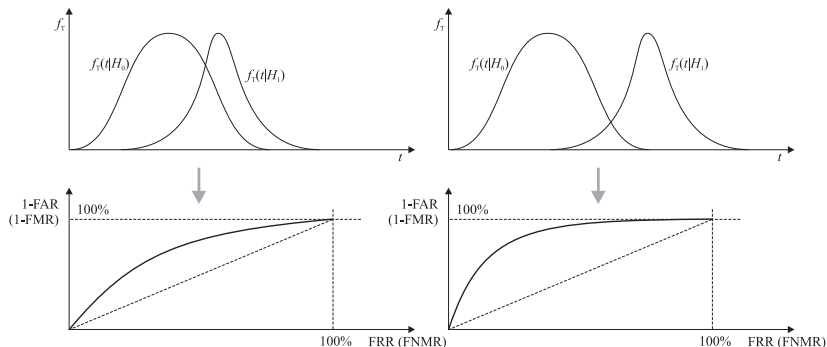




# Ocena systemów weryfikacji

## Krzywa ROC (ang. Receiver Operating Characteristic)

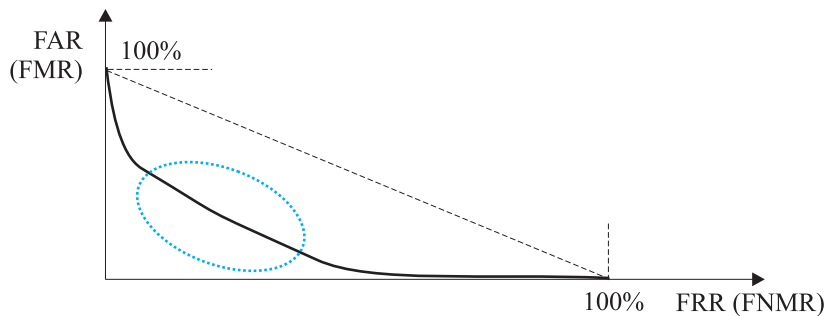
Pole pod krzywą ROC to **miara jakości klasyfikacji** (w szczególności 1.0 dla idealnego klasyfikatora, 0.5 dla klasyfikatora losowego)



# Ocena systemów weryfikacji

## Krzywa DET (ang. Detection Error Tradeoff)

Wartości estymatorów błędów wyrażone w skali **odwrotnej dystrybuanty rozkładu normalnego** (ang. normal deviate scale)



# Ocena systemów weryfikacji

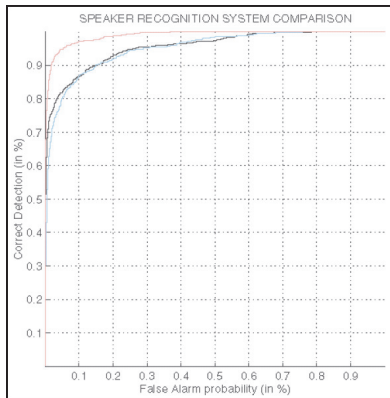
## Własności DET

1. Dla rozkładów normalnych: proste na diagramie DET
2. Nachylenie prostych: iloraz wariancji rozkładów
3. Odległość prostych: różnica wartości średnich rozkładów
4. Dla klasyfikatora losowego zmiennych losowych o rozkładzie normalnym (o tych samych parametrach) DET jest prostą  $FAR=1-FRR$  (lub  $FMR=1-FNMR$ )

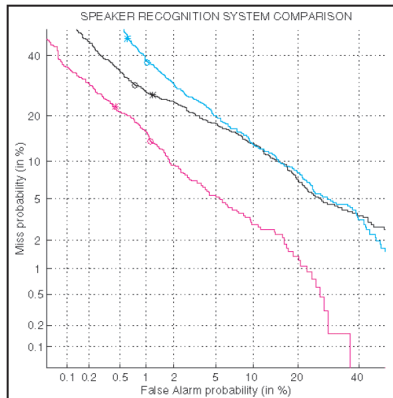
# Ocena systemów weryfikacji

## Porównanie ROC i DET dla tych samych ocen (przykład)

### ROC



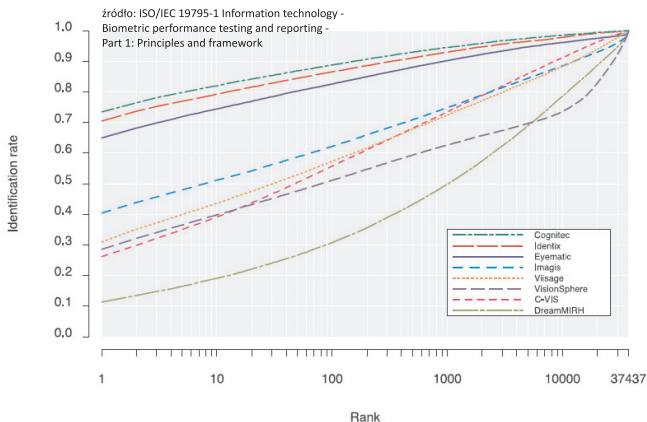
### DET



źródło: A. Martin, G. Doddington, T. Kamm, M. Ordowski, M. Przybocki, „The DET curve in assessment of detection task performance”, EuroSpeech 1997, Proceedings Volume #4, pp. 1895-1898, NIST 1997

## Ocena systemów identyfikacji

Skumulowana charakterystyka dopasowania CMC (ang. Cumulative Match Characteristic): **częstość identyfikacji rzędu  $r$**  w funkcji  $r$



## Ocena systemów identyfikacji

1. **Systemy zamknięte:** wszyscy potencjalni użytkownicy zostali zarejestrowani
  - **częstość identyfikacji rzędu  $r$ :** stosunek liczby transakcji, w których zarejestrowany użytkownik znalazł się na liście najlepszych  $r$  kandydatów do liczby transakcji
2. **Systemy otwarte:** nie wszyscy potencjalni użytkownicy zostali zarejestrowani
  - **częstość identyfikacji rzędu  $r$  oraz dodatkowo:**
  - **częstość fałszywie negatywnej identyfikacji** (ang. False Negative Identification Rate – FNIR): proporcja transakcji identyfikacyjnych użytkowników **zarejestrowanych**, dla których właściwego identyfikatora **nie ma na liście** kandydatów
  - **częstość fałszywie pozytywnej identyfikacji** (ang. False Positive Identification Rate – FPIR): proporcja transakcji identyfikacyjnych użytkowników **niezarejestrowanych**, dla których lista kandydatów **nie jest pusta**

## Wykład 10: Statystyczna ocena systemów biometrycznych

Problem ewaluacji w biometrii

Modelowanie niepewności

Ocena procesów pozyskiwania próbek i rejestracji

Ocena zgodności próbek

Ocena systemów biometrycznych

**Ocena rozrzutu parametrów**

Wybrane uwagi praktyczne

# Estymacja przedziałowa

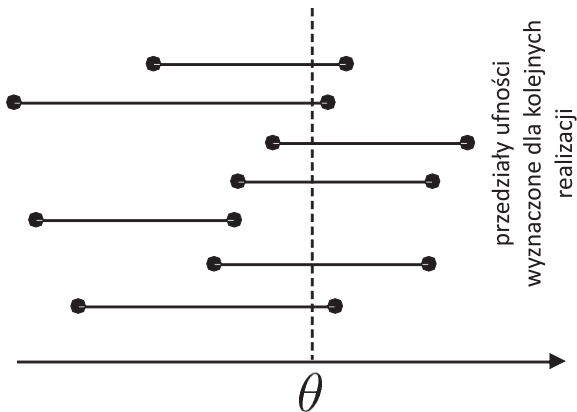
## Pojęcie przedziału ufności

1. Estymator przedziałowy wyznaczają dwie zmienne losowe
2. Przedziały ufności
  - otrzymane wartości estymatora przedziałowego na podstawie obserwacji próby
  - końce  $\theta_1$ ,  $\theta_2$  przedziału ufności dla nieznanego parametru  $\theta$  są funkcjami próby losowej, tzn.  $\theta_1 = g_1(X_1, \dots, X_n)$ ,  $\theta_2 = g_2(X_1, \dots, X_n)$  i **nie zależą** od  $\theta$
  - prawdopodobieństwo pokrycia przez przedział  $\langle \theta_1, \theta_2 \rangle$  szacowanego parametru  $\theta$  wynosi  $1 - \alpha$  ( $\alpha$  to poziom istotności)
  - **uwaga:** **nie jest prawdą**, że  $\theta$  leży w  $\langle \hat{\theta}_1, \hat{\theta}_2 \rangle$  obliczonym dla jednej obserwacji próby losowej z prawdopodobieństwem  $1 - \alpha$



# Estymacja przedziałowa

## Pojęcie przedziału ufności



# Estymacja przedziałowa w biometrii

## Maksymalne prawdopodobieństwo popełnienia błędu

1. Niezależne wyniki porównań biometrycznych, przy niezmiennym, danym prawdopodobieństwie porażki  $p_e$  można traktować jak próby Bernoulliego
2. Prawdopodobieństwo zaobserwowania do  $N_e$  błędów w próbie  $N_c$  elementowej

$$\mathcal{P}(\xi \leq N_e) = \sum_{n=0}^{N_e} \frac{N_c!}{n!(N_c - n)!} p_e^n (1 - p_e)^{N_c - n}$$

3. Granica  $p_e^+$  przedziału ufności oraz założony poziom istotności  $\alpha$  wiąże więc równanie

$$\sum_{n=0}^{N_e} \frac{N_c!}{n!(N_c - n)!} (p_e^+)^n (1 - p_e^+)^{N_c - n} = \alpha$$

## Estymacja przedziałowa w biometrii

### Maksymalne prawdopodobieństwo popełnienia błędu

4. Nie zaobserwowaliśmy błędu (z próby). Jaki jest najniższy, gwarantowany statystycznie dla założonego  $\alpha$ , błąd dla systemu biometrycznego? Przyjmując  $N_e = 0$  otrzymujemy:

$$(1 - p_e^+)^{N_c} = \alpha \quad \text{czyli} \quad N_c \ln(1 - p_e^+) = \ln(\alpha)$$

Przyjmując przykładowo 95% poziom ufności ( $\alpha = 0.05$ ), zakładając, że

$$\ln(0.05) \equiv -3$$

oraz, że dla małych  $p_e^+$

$$\ln(1 - p_e^+) \equiv -p_e^+$$

otrzymujemy

$$p_e^+ = \frac{3}{N_c} \quad (\text{tzw. "reguła trzech"})$$

## Estymacja przedziałowa w biometrii

### Maksymalne prawdopodobieństwo popełnienia błędu

5. Dla porównań **zależnych**, bądź w sytuacji **zmiennego  $p_e$**  dla populacji testowej, **nie możemy** traktować wyników porównań jako prób Bernoulliego
6. Kiedy wyniki porównań w biometrii są zależne?
  - te same obiekty uczestniczą w wielu porównaniach podczas szacowania wartości błędu FA (np. porównania krzyżowe, ang. cross-comparisons)
  - wykonujemy wielokrotne próby dla tego samego użytkownika podczas szacowania błędu FR (np. użytkownik nie podejmuje prób po udanej weryfikacji, podejmuje natomiast kolejną próbę po odrzuceniu go przez system)

## Wykład 10: Statystyczna ocena systemów biometrycznych

Problem ewaluacji w biometrii

Modelowanie niepewności

Ocena procesów pozyskiwania próbek i rejestracji

Ocena zgodności próbek

Ocena systemów biometrycznych

Ocena rozrzutu parametrów

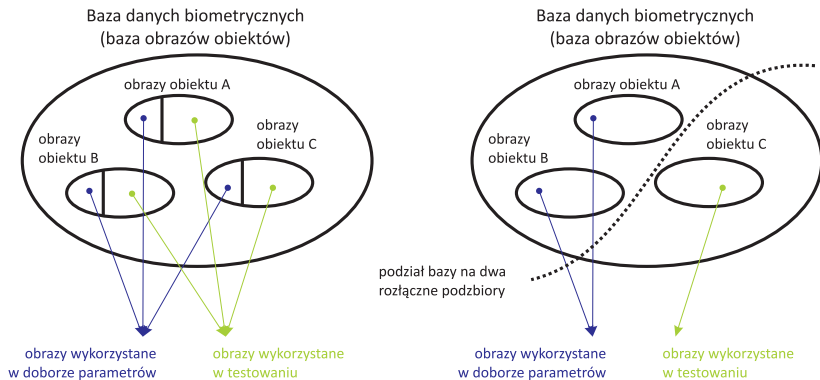
Wybrane uwagi praktyczne

## Zasady wykorzystania biometrycznych baz danych

1. Podział bazy danych na rozłączne zbiory:
  - **estymacyjny**: do estymacji parametrów metod
  - **testowy**: do przeprowadzania ewaluacji
2. Istnieje wiele możliwych podziałów bazy
  - staramy się, aby wynikowe porównania były **niezależne** → wymaga “dużych” zbiorów
  - przy “małych” zbiorach lepiej analizować **więcej zmiennych zależnych** niż “zbyt mało” niezależnych
  - właściwe metody statystyczne dla “zbyt małych” zbiorów, np. bootstrap

# Zasady wykorzystania biometrycznych baz danych

## Przykłady podziału bazy danych biometrycznych



## Projektowanie rozpoznawania biometrycznego

### 1. EER nie najważniejsze

- punkt pracy wykorzystywany często w ocenie laboratoryjnej komponentów rozwiązania, rzadko w ocenie systemów komercyjnych
- najczęściej w testach wyliczana wartość FRR przy założonym poziomie FAR (zwykle niskim, np.  $FAR=0.001$ )

### 2. Różne punkty pracy systemu w zależności od zastosowania

- systemy o wysokim stopniu bezpieczeństwa: nacisk na małe wartości FAR (kosztem FRR)
- przeciwnie dla systemu o dużym stopniu wygody użytkownika: chęć zmniejszania FRR (kosztem FAR)
- metody testujące żywotność obiektów: małe FRR (dowolne FAR) ze względu na niemożność modelowania rozkładów dla fałszerstw zaawansowanych, oraz większe (z punktu widzenia odbioru technologii) konsekwencje zwiększania FRR niż ewentualnej próby fałszerstwa (dla niektórych systemów mało prawdopodobnego)



## Przykładowe pytanie egzaminacyjne

Testujemy dwa algorytmy rozpoznawania tęczówki wykorzystujące algorytm Daugmana: A i B. Dla systemu A otrzymaliśmy następujące wyniki porównań tych samych tęczówek: 0.1, 0.2, 0.3, oraz różnych tęczówek: 0.4, 0.5, 0.6. Dla systemu B zostały odnotowane następujące wyniki porównań tych samych tęczówek: 0.1, 0.2, 0.4, oraz różnych tęczówek: 0.3, 0.5, 0.6. Dla którego z systemów pole pod krzywą ROC jest większe? Uzasadnij odpowiedź.