

Over het nut van herkansen

Robert Zwitser MSc

Dr. Anton Béguin

robert.zwitser@cito.nl

Waarom herkansen?

- Mogelijkheid om bij te leren
- De rol van willekeur in de toetsscore
 - Keuze van de toetsvragen uit het domein
 - Omstandigheden bij de afname (tijdstip, rumoer, etc.)
 - De keuze van de beoordelaar
 - ...

Waarom herkansen?

Bij een replicatie van de toetsafname kan er hierdoor een andere score ontstaan.

De invloed van deze willekeur noemen we **meetfout**.

Bij beslissingen o.b.v. toetsscores kan bij een replicatie van de toetsafname een andere beslissing worden genomen.

Waarom herkansen?

Herkansing is ook voor **onterecht** gezakte kandidaten -> nieuwe kans

- Eerlijkheid
- Betrouwbaarheid

		Werkelijkheid	
		Gezakt	Geslaagd
Beslissing	Gezakt	Terecht	Onterecht
	Geslaagd	Onterecht	Terecht

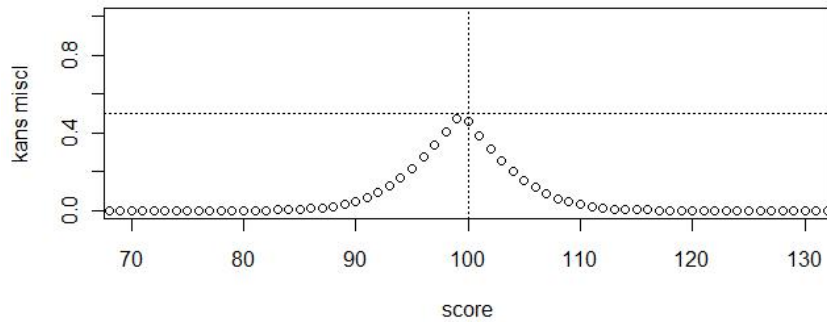


Herkansing en betrouwbaarheid

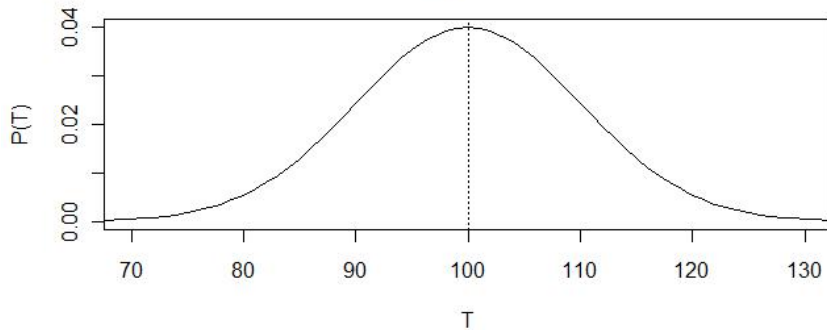
- Wie herkansen?
 - Alle gezakte kandidaten
 - Alle kandidaten met een kans op een andere uitkomst
- Welke score telt?
 - De hoogste score
 - Beide scores samen (toetsverlenging)
- Positie cesuur

Voorbeeld één toets

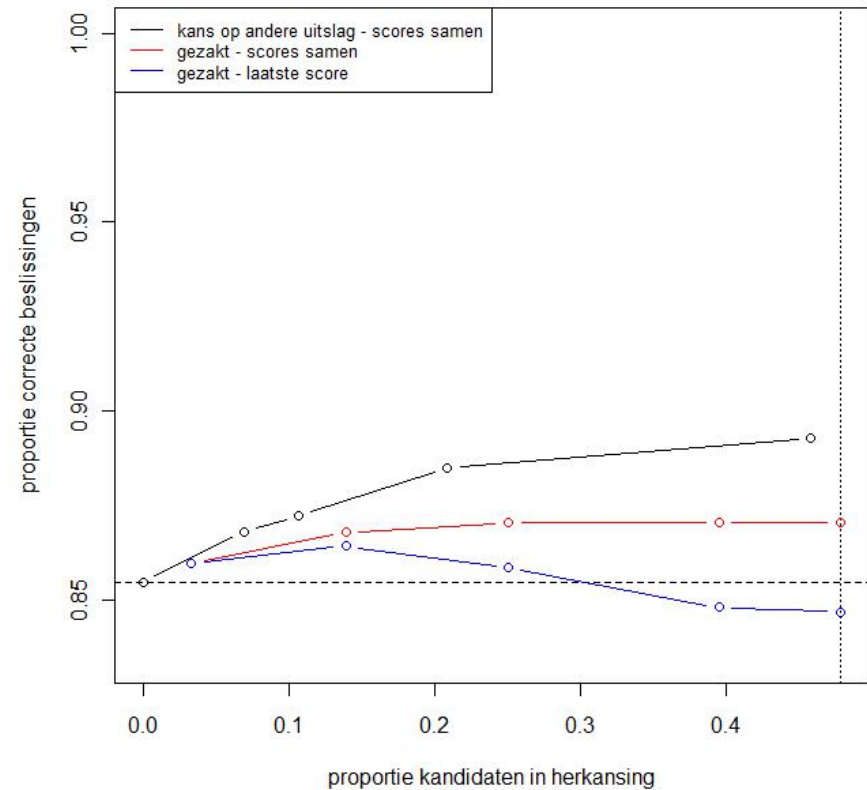
kans op misclassificatie



verdeling ware scores

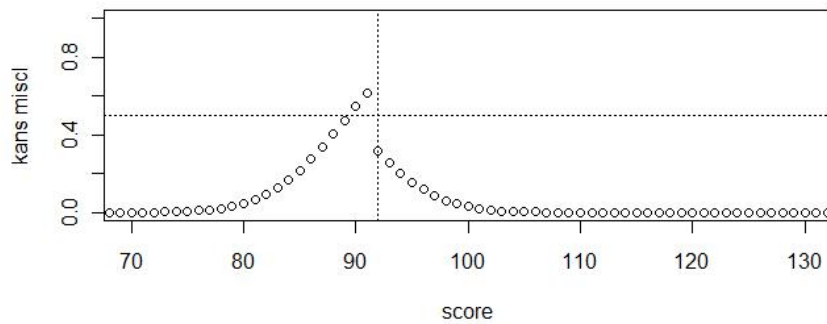


rho: 0.8, cesuur: 50%

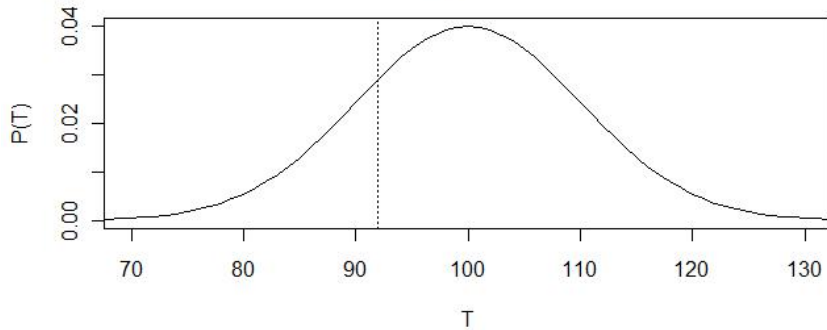


Voorbeeld één toets

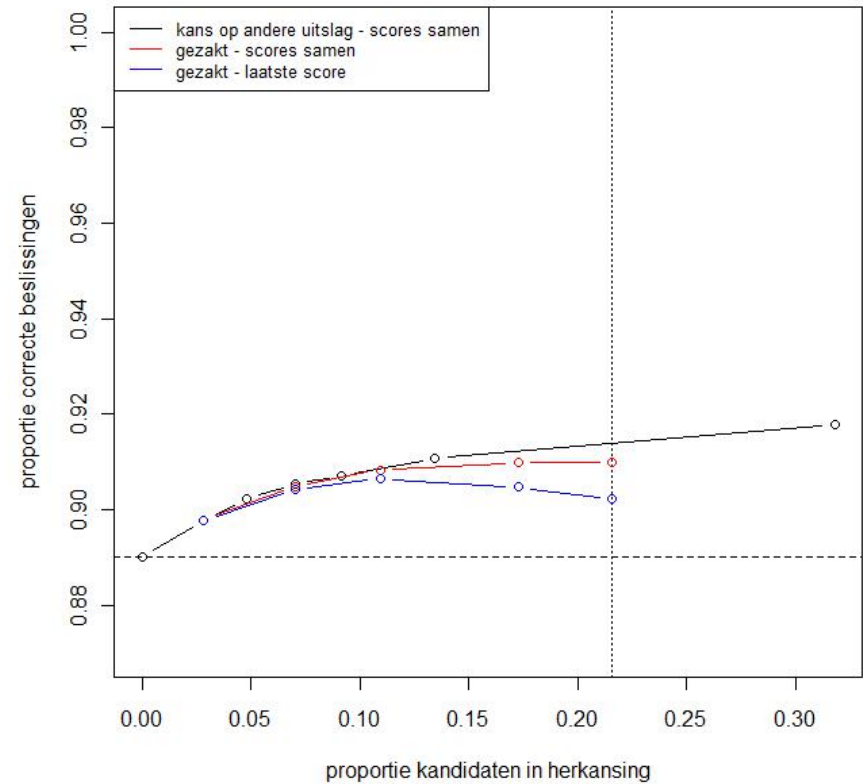
kans op misclassificatie



verdeling ware scores

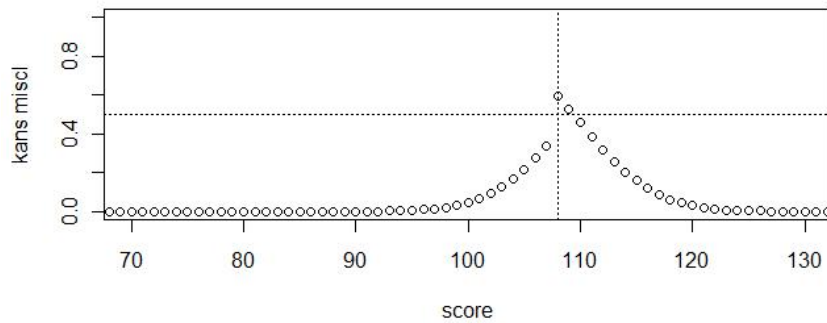


rho: 0.8, cesuur: 20%

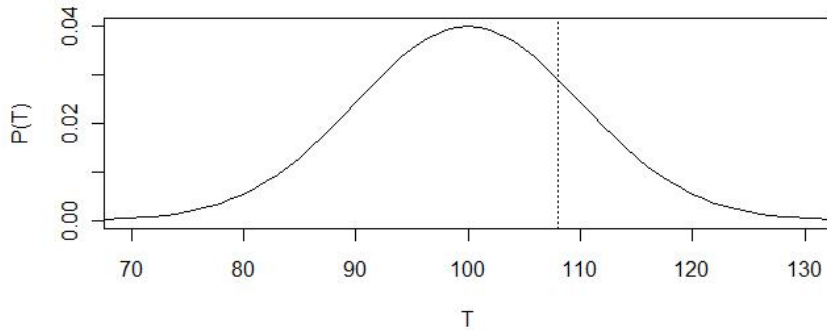


Voorbeeld één toets

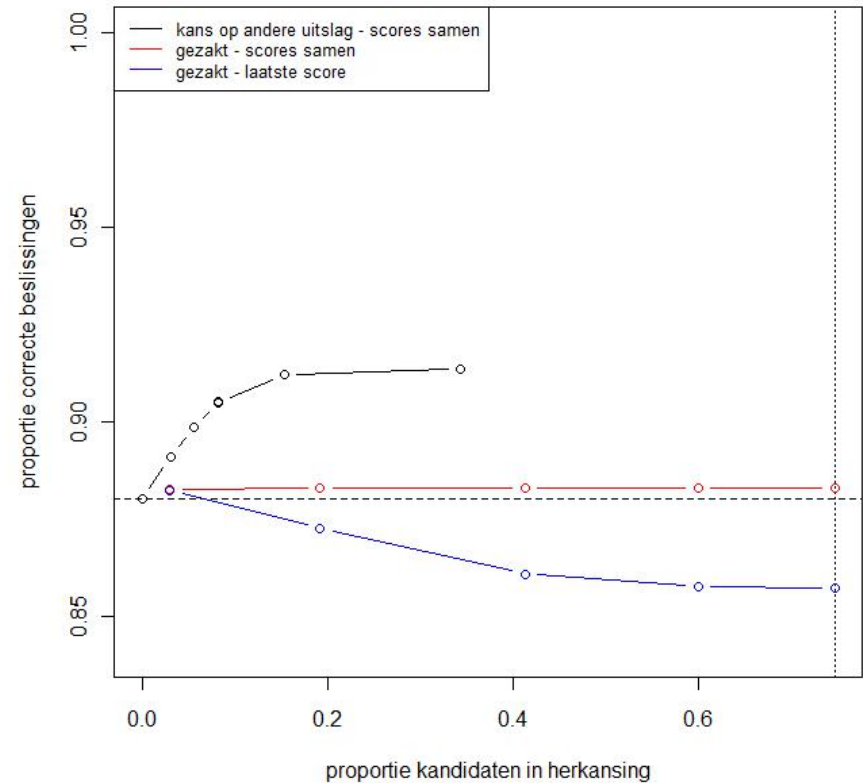
kans op misclassificatie



verdeling ware scores

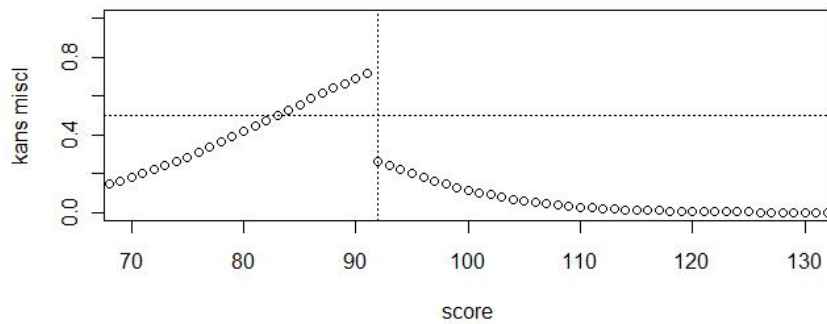


rho: 0.8, cesuur: 80%

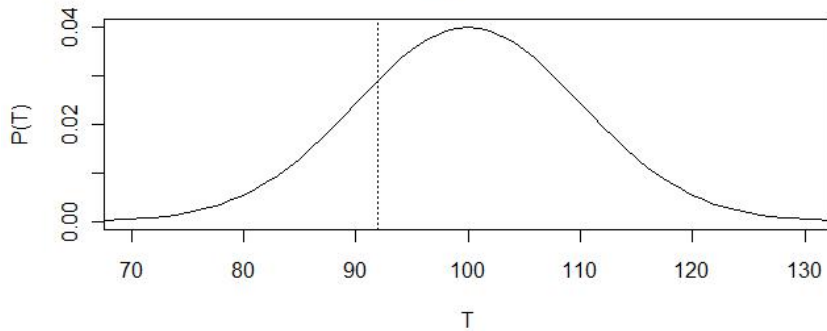


Voorbeeld één toets

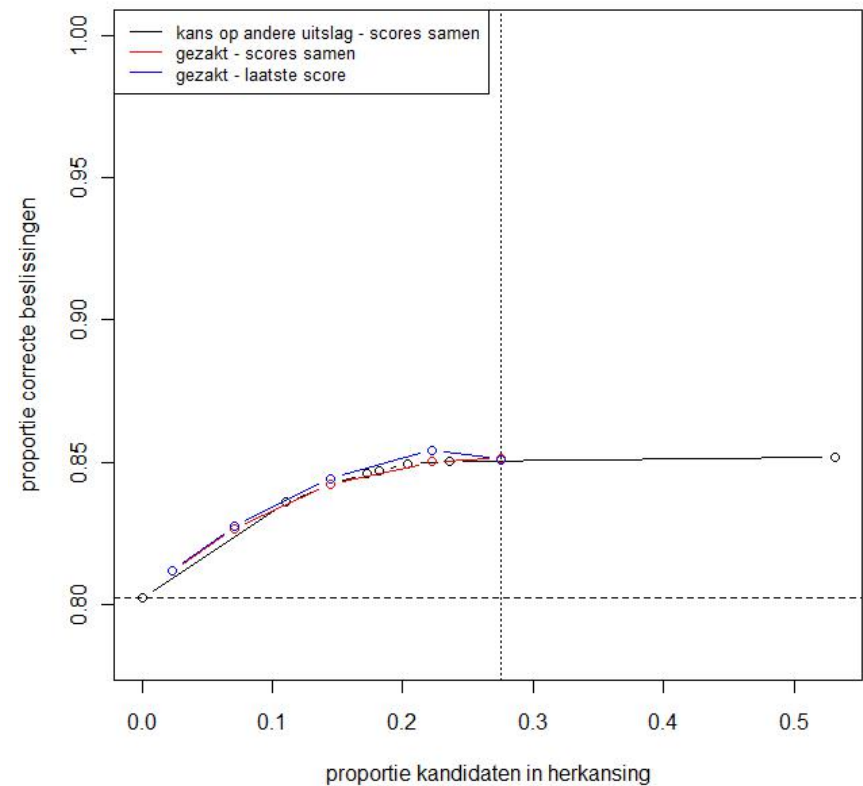
kans op misclassificatie



verdeling ware scores



rho: 0.5, cesuur: 20%



Praktijkvoorbeeld examens

- Meerdere scores -> één beslissing
- Correlatie tussen de scores
- Compensatiemogelijkheid

Betrouwbaarheid:

Percentage kandidaten dat bij een replicatie van de toetsafname op gelijke wijze geclassificeerd zou worden (*Standards, AERA, APA & NCME, 1999*).

Voorbeeld examens

- E&M profiel Havo 2010
 - NE, EN, DU, EC, GS, WI A, M&O
- Steekproef 500 leerlingen
- 7 dimensioneel IRT model
 - MCMC algoritme
 - Plausible values
- Replicaties scores / cijfers / uitslag

Voorbeeld examens

Eerste afname:

		Replicaties (20x)		
		Gezakt	Geslaagd	Totaal
Werkelijke cijfers	Gezakt	98.2	26.8	125 (25)
	Geslaagd	12.6	362.4	375 (75)
	Totaal	110.8 (22)	389.2 (78)	

Betrouwbaarheid: 0.92

Voorbeeld examens

Leerl. met kans op andere uitslag: 104 / 500

Vak	SE	CE	EIND	% anders	Gem. repl.	Sd. repl
NE	6.7	5.5	6	0	5.99	0.21
EN	5.9	3.7	5	0	5.03	0.35
DU	6.2	6.3	6	0	6.09	0.21
EC	6.2	6.1	6	2/20	5.98	0.35
GS	5.8	5.2	6	7/20	5.55	0.20
WIA	6.7	7.5	7	0	6.91	0.23
M&O	6.4	5.0	6	0	5.92	0.19



zeker weten

Voorbeeld examens

- Leerlingen met kans op andere uitslag
- Resultaten samen nemen

		Replicaties		
		Gezakt	Geslaagd	Totaal
Werkelijke cijfers	Gezakt	95.50	28.50	124
	Geslaagd	13.55	362.45	376
	Totaal	109.05	390.95	

Betrouwbaarheid: 0.92

Voorbeeld examens

- Gezakte leerlingen
- Resultaten samen nemen

		Replicaties		
		Gezakt	Geslaagd	Totaal
Werkelijke cijfers	Gezakt	97.45	26.55	124
	Geslaagd	12.15	363.85	376
	Totaal	109.60	390.40	

Betrouwbaarheid: 0.92

Voorbeeld examens

- Gezakte leerlingen
- Hoogste score

		Replicaties		
		Gezakt	Geslaagd	Totaal
Werkelijke cijfers	Gezakt	85.7	22.3	108
	Geslaagd	18.3	373.7	392
	Totaal	104	396	

Betrouwbaarheid: 0.92

Conclusies / discussie

- Toetsverlenging / herkansing kan een positief effect hebben op de betrouwbaarheid (wisten we al...)
- Welke scores een hoge kans op misclassificatie hebben, wordt zowel bepaald door de positie van de cesuur als de betrouwbaarheid van de toets.

Conclusies / discussie

- Eén herkansing heeft bij de eindexamens nauwelijks effect op de betrouwbaarheid van de beslissing.
- Rol compensatie / effect bij conjunctieve regels
- Meerdere herkansingen -> adaptief toetsen

Vragen?

robert.zwitser@cito.nl