

# VQR e valutazione interna dei Dipartimenti

## *Sommario:*

- *Perché il Gruppo di lavoro CRUI ha proposto all'ANVUR un nuovo modo per calcolare indicatori della qualità della ricerca dei dipartimenti (secondo VQR)*
- *L'approccio proposto per la erogazione di risorse non dipendenti dal costo della ricerca*
- *Le proprietà dell'indicatore "Voto Standardizzato"*
- *I risultati dell'indicatore sul sistema universitario*
- *Il Voto Standardizzato e il suo possibile utilizzo per l'erogazione delle risorse non dipendenti dal costo della ricerca*

## L'esercizio VQR e i dati raccolti

*L'esercizio VQR: uno sforzo notevole per tutto il sistema degli Atenei e degli Enti di Ricerca.*

*Disponibilità di una mole impressionante di dati.*

*Questi dati devono essere utilizzati nella loro completezza e al meglio.*

*La proposta del Gruppo di lavoro CRUI all'ANVUR: utilizzo pieno della informazione, sfruttando al massimo il dettaglio disponibile (cioè di SSD)*

## Gli indicatori VQR di Dipartimento

*L'indicatore IRFD ANVUR non è particolarmente adatto per distribuire risorse non dipendenti dal costo della ricerca:*

- *Dipartimenti fortemente non-omogenei in termini di afferenza alle varie Aree CUN;*
- *Presenza di voci non applicabili a tutte le Aree e/o SSD;*
- *Pesi riferiti al costo della ricerca nell'indicatore complessivo di dipartimento (IRFD), oltretutto basato sulla combinazione delle valutazioni di Area e non di SSD;*

## Dipartimenti e indicatori VQR

*Ci sarebbe un'altra domanda fondamentale: come si colloca, nel panorama internazionale, un certo SSD nazionale? Nel rapporto finale ANVUR, Parte Terza viene toccato il punto (per le Aree CUN "bibliometriche") e forse esso merita, prima o poi, un approfondimento.*

*In questo documento tale aspetto non è trattato; ciò non significa che non sia importante, anzi. Il prossimo esercizio VQR dovrebbe affrontare con decisione questo tema.*

*Di ciò in particolare andrebbe tenuto conto per erogare risorse legate alle attività di ricerca.*

## Un indicatore non dipendente dal costo della ricerca

*Per distribuire ai dipartimenti risorse quali i punti organico sarebbe opportuno:*

- 1. Utilizzare esclusivamente la valutazione dei prodotti della ricerca a livello di SSD (informazione più dettagliata)*
- 2. Non utilizzare la capacità di attrarre risorse, di compiere trasferimento tecnologico, ovvero di tutto ciò che non sia applicabile a tutti i SSD;*
- 3. E poi, soprattutto....*

## Un indicatore non dipendente dal costo della ricerca

... individuare un metodo per confrontare fra di loro e ordinare secondo la qualità della ricerca dipartimenti aventi membri appartenenti a SSD molto diversi e con numerosità diverse,

tenendo presente che –ovviamente-- ogni Ateneo è interessato ad avere i migliori professori e ricercatori che afferiscono ai vari SSD presenti al suo interno.

## Utilizzare il criterio del percentile superiore (top %)

- *L'approccio suggerito si ispira alla classificazione che le università statunitensi ci richiedono talvolta: in quale "top%" dei nostri voti si colloca lo studente che ha fatto domanda di iscriversi presso di loro?*
- *Questa classificazione rimuove la dipendenza esplicita dalla modalità di valutare e per farla occorre solo avere la distribuzione dei voti assegnati*

## Classificare votazioni di SSD con i percentili superiori

- Con questo stesso metodo possiamo confrontare le valutazioni VQR in due SSD completamente differenti. A seconda del percentile della distribuzione delle votazioni nazionali dei rispettivi SSD nel quale esse si collocano, decidiamo quale delle due valutazioni è migliore o peggiore
- Per essere applicato, questo criterio richiede esclusivamente di conoscere le distribuzioni dei voti assegnati dalla VQR ai vari membri dei due SSD considerati (che sono disponibili sul sito ANVUR-VQR).



## Come classificare i dipartimenti?

*E per confrontare dipartimenti, come fare?*

- *Occorre definire un indicatore “collettivo” di dipartimento (fattibile), crescente con i voti dei membri;*
- *Purtroppo è impossibile valutare il percentile in cui esso si colloca;*
- *Per farlo occorrerebbe disporre di un numero molto ampio di dipartimenti composti in termini di SSD esattamente come il nostro. Ovviamente ciò non si verifica mai e ogni dipartimento costituisce un caso a sé stante.*

## Come classificare i dipartimenti?

*Possiamo però ragionare così:*

- *Per ogni dipartimento reale (DR) dell'ateneo immaginiamo di avere un Dipartimento Virtuale o "Dipartimento Specchio" ad esso associato (DVA nel seguito): questo è un dipartimento ipotetico composto da membri (fittizi) che afferiscono ai vari SSD come nel nostro dipartimento reale DR, con la stessa numerosità.*
- *Calcoliamo tutti i valori che l'indicatore può assumere nel DVA, basandoci sulle valutazioni nazionali dei rispettivi SSD;*

*Facciamo un esempio concreto di un possibile indicatore*

## Come classificare i dipartimenti?

*Come indicatore utilizziamo l'Indicatore di Dipartimento Virtuale Associato, definito come segue:*

$$IDVA = \frac{1}{N} \sum_{ssd=1}^{NSSD} R_{ssd} \cdot n_{ssd}$$

$$R_{ssd} = \frac{I_{ssd}}{\langle I_{ssd} \rangle}$$

*$R_{ssd}$  è il cosiddetto "terzo indicatore" della VQR; rapporto fra il voto medio dei prodotti di ricerca nel SSD del dipartimento rispetto a quello medio nazionale.*

*IDVA rappresenta il valore di R mediato sui SSD presenti del dipartimento. L'indicatore compensa (solo) per i differenti valori medi diversi nei vari SSD. Come vedremo, ne possiamo trovare di migliori. Ora serve per illustrare il metodo.*

# Un esempio di calcolo di *IDVA*

$$IDVA = \frac{1}{N} \sum_{SSS=1}^{NSSD} R_{ssd} \cdot n_{ssd}$$

SSD	prodotti	Voto SSD nel Dip	Voto Nazionale	R	R*n
VET/01	90	0.75	0.73	1.027	92.466
VET/03	120	0.81	0.84	0.964	115.714
AGR/19	9	0.56	0.58	0.966	8.690
Totale	219				216.870

$$IDVA = (92.466+115.714+8.690)/219 = 0.990$$

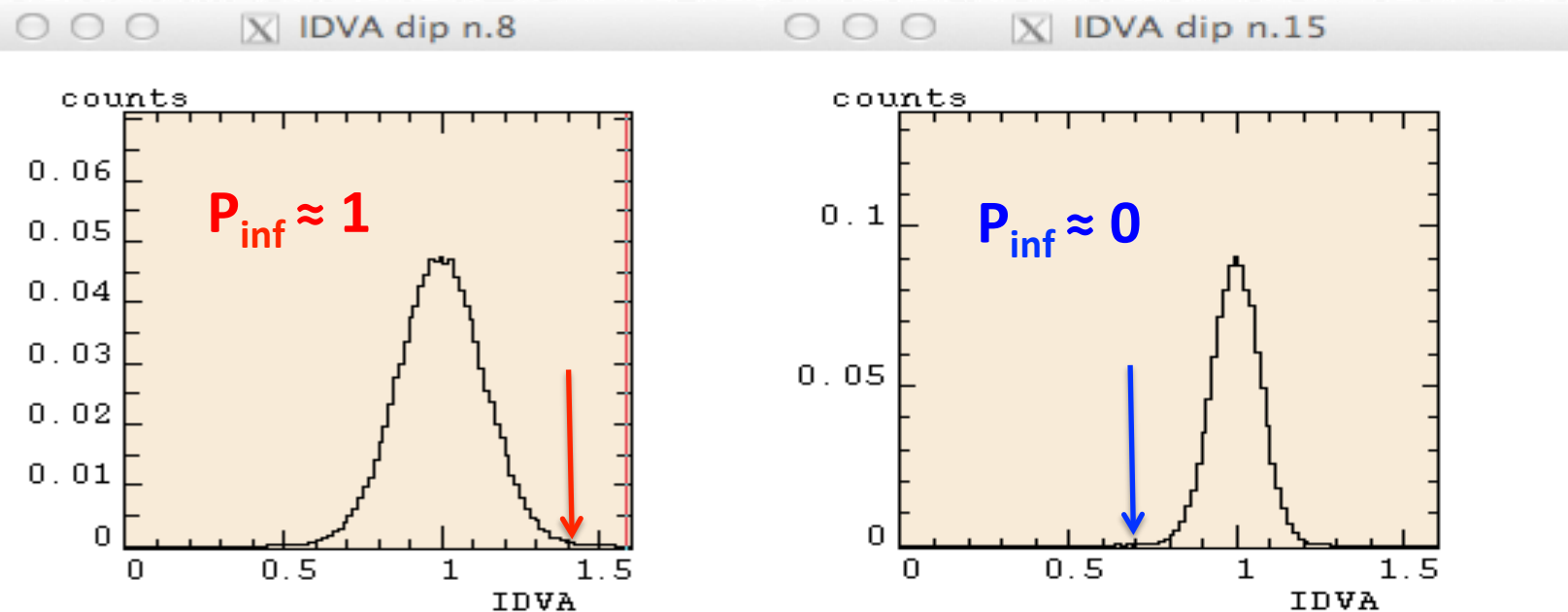
## Come classificare i dipartimenti?

- *l'indicatore  $IDVA_d$  assume per il DR il valore  $IDVA^*_d$  (notare \*)*
- *Nel DVA, l'indicatore può assumere numerosissimi valori  $IDVA_d$  (niente \*), tutti quelli ottenibili dalle combinazioni dei voti nazionali dei SSD; il calcolo può essere effettuato in modo efficiente con metodi di simulazione (Metodo Montecarlo)*
- *Contando il numero di casi in cui il valore dell'indicatore  $IDVA_d$  è inferiore a  $IDVA^*_d$  si determina la probabilità  $P_{inf}(IDVA^*_d)$  di ottenere un  $IDVA_d$  nel DVA con un valore dell'indicatore inferiore a quello di DR.*
- *Sinteticamente possiamo dire che se  $P_{inf}(IDVA^*_d)$  è prossimo a 1 (uno) il DR è molto buono; se viceversa  $P_{inf}(IDVA^*_d)$  è prossima 0, il DR è molto modesto.*

# Il confronto con il Dipartimento Virtuale Associato

Si confronta il valore di  $IDVA^*_d$  sul DR con la distribuzione di tutti i valori possibili che  $IDVA_d$  può assumere nel DVA (Simulazioni Montecarlo).

Un dipartimento buono a sinistra e un dipartimento modesto a destra: le frecce indicano i valori di  $IDVA^*_d$ , cioè i valori dell'indicatore sui DR, le curve tutti i valori possibili di  $IDVA_d$  sui DVA.



Nel caso di sinistra praticamente tutti i valori di  $IDVA_d$  del DVA sono inferiori a quello reale; l'opposto per il dipartimento di destra. Le due "campane" sono diverse.

## Il confronto con il Dipartimento Virtuale Associato

*La conoscenza di  $P_{inf}(IDVA^*_d)$  permette di definire un metodo operativo per confrontare la qualità VQR di qualunque insieme di dipartimenti fra di loro.*

*Migliori saranno quelli che sono caratterizzati da valori di  $P_{inf}(IDVA^*_d)$  maggiori*

### *Problemi di applicabilità:*

- Montecarlo su 850 dipartimenti nazionali – un po' laborioso, possibile, ma poco trasparente (solo pochi utilizzano questo metodo)*
- Occorre cercare di correggere le variabilità valutative dei GEV. Da questo punto di vista il terzo indicatore (R e quindi anche IDVA che è un "R mediato sui SSD") lascia a desiderare.*

*La soluzione proposta per entrambi i problemi: il Voto Standardizzato.*

*[http://www.anvur.org/index.php?option=com\\_content&view=article&id=609&Itemid=563&lang=it9](http://www.anvur.org/index.php?option=com_content&view=article&id=609&Itemid=563&lang=it9)*

## Il Voto Standardizzato e il calcolo di $P_{inf}$

*Dal campione nazionale completo delle valutazioni del SSD  $s$  estraiamo la votazione media  $\langle I_s \rangle$  dei soggetti valutati e la deviazione standard campionaria  $\sigma_s$  che misura la “larghezza della distribuzione”.*

*Esaminando i vari SSD nazionali, si trova che i valori di  $\langle I_s \rangle$  e  $\sigma_s$  cambiano fortemente da SSD a SSD.*

*Spesso piccoli valori di  $\langle I_s \rangle$  si accompagnano a grandi valori di  $\sigma_s$  e viceversa.*

*Definiamo allora, per ciascun membro  $l$  del dipartimento  $d$  afferente al SSD  $s$  il suo voto standardizzato  $u_{l,s,d}$  :*

$$u_{l,s,d} = \frac{v_{l,s,d} - \langle I_s \rangle}{\sigma_s}$$

*In maniera un po' semplificata:*

*il voto standardizzato compensa per i diversi voti medi (differenza fra voto ricevuto e voto medio); compensa parzialmente per le diverse larghezze delle distribuzioni dei voti nazionali (rapporto con la deviazione standard campionaria del SSD).*



## Il Voto Standardizzato di dipartimento e $P_{inf}$

*Mentre le distribuzioni delle valutazioni dei vari SSD hanno diversi valori medi  $\langle I_s \rangle$  e diverse deviazioni standard  $\sigma_s$ , le corrispondenti distribuzioni del VS, qualunque sia il SSD esaminato, hanno tutte lo stesso valor medio (0) e la stessa deviazione standard (1) -- per gli esperti: solo i momenti di ordine superiore sono, in genere, differenti.*

*Altro vantaggio (matematicamente correlato al precedente): partendo dai VS di singolo membro è possibile costruire, semplicemente sommandoli e normalizzandoli, l'indicatore Voto Standardizzato di Dipartimento  $U_d^*$  che gode di alcune importanti e utili proprietà.*

## Il Voto Standardizzato di dipartimento e $P_{inf}$

Il VS di dipartimento  $U_d$  è definito come la media dei VS di singolo membro, pesata con il numero di prodotti che i vari soggetti valutati dovevano presentare:

$$U_d = \frac{1}{\sqrt{\sum_{s=1}^{NSSD} \sum_{l=1}^{N_{s,d}} np_{l,s,d}^2}} \sum_{s=1}^{NSSD} \sum_{l=1}^{N_{s,d}} u_{l,s,d} \cdot np_{l,s,d}$$

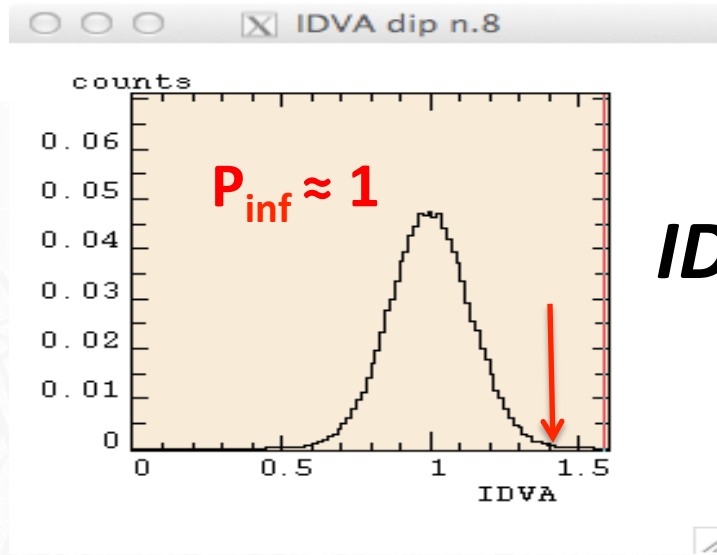
Il denominatore assicura che l'insieme dei valori  $\{U_d\}$  calcolati per il DVA abbia valore medio pari a 0 e deviazione standard pari a 1. Assicura inoltre che l'indicatore sia depurato dall'effetto dimensione.

Si può dimostrare che sussistono tutte le condizioni matematiche che assicurano che la distribuzione dei valori di  $\{U_d\}$  sul DVA, qualunque sia il DR che stiamo esaminando, sia sempre la stessa: una gaussiana standardizzata.

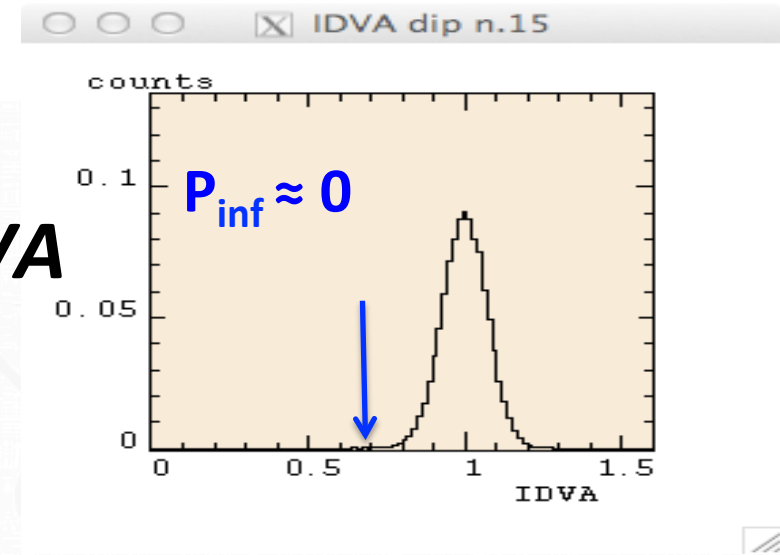
L'uguaglianza di tutte le distribuzioni di  $\{U_d\}$  semplifica enormemente il calcolo, altrimenti complesso e laborioso (ricorso al Montecarlo), di  $P_{inf}(U_d^*)$ : il calcolo è lo stesso per tutti i DR.

# Il Voto Standardizzato di dipartimento e $P_{inf}$

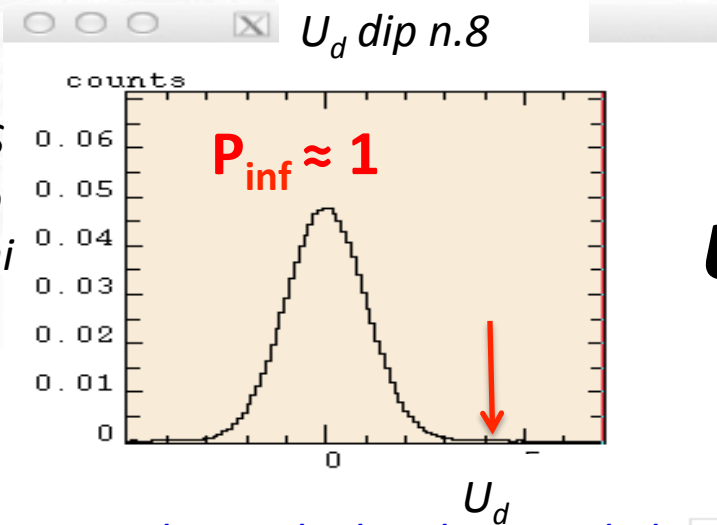
Calcolo di  $P_{inf}$  utilizzando  $IDVA_d$ . Distribuzioni diverse per dipartimenti diversi



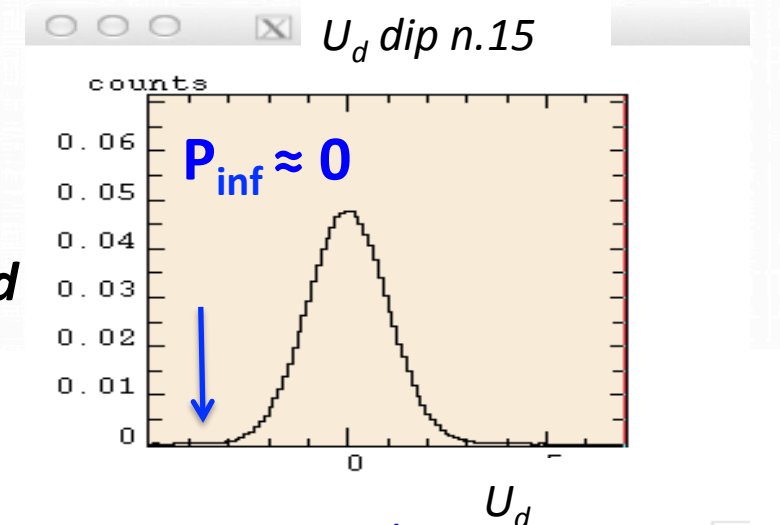
**IDVA**



Calcolo di  $P_{inf}$  utilizzando il VS di dipartimento  $U_d$ . Distribuzioni uguali per dipartimenti diversi.

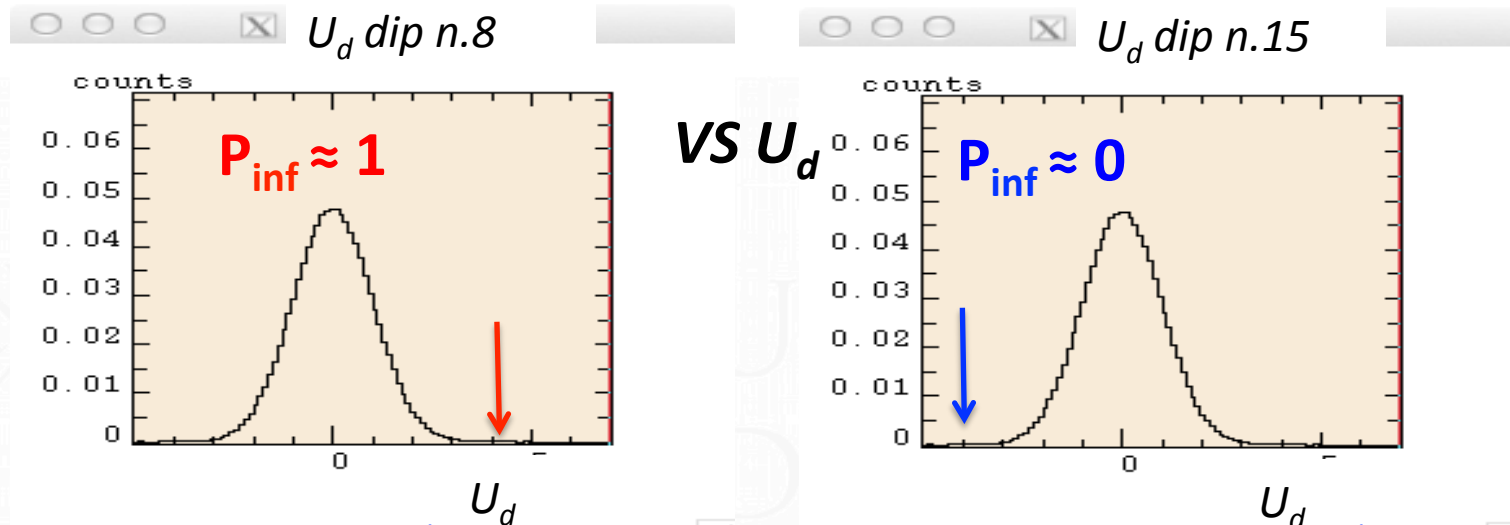


**$U_d$**



Notare che ora le distribuzioni di destra e sinistra sono uguali

# Il Voto Standardizzato di dipartimento e $P_{inf}$



Calcolare diretto di  $P_{inf}(U_d^*)$  per qualunque dipartimento, noto  $U_d^*$ : notevole semplificazione. Nessuna necessità di ricorrere agli adepti del Montecarlo. Basta fare un calcolo da foglio elettronico:  $P_{inf}(U_d^*) = \text{DISTRIB.NORM.}(U_d^*, 0, 1, 1)$ .

Uguaglianza di tutte le distribuzioni di  $U_d$  sui DVA garantita dal Teorema del Limite Centrale (TLC). Rigorosamente ciò vale solo nel caso di un numero infinito di membri del dipartimento.

Per scrupolo è stato verificato, studiando i circa 900 dipartimenti nazionali, che la massima differenza fra il calcolo che assume la validità del TLC e il calcolo da simulazione Montecarlo ammonta al massimo al 3.5% (un Dipartimento del Politecnico di Milano).

# Il Voto Standardizzato di dipartimento e $P_{inf}$

## *Limiti noti del VS:*

- *La compensazione non perfetta per i diversi voti medi e le diverse larghezze;*
- *Infatti i valori possibili (accessibili) del VS non sono gli stessi per tutti i SSD;*
- *In pratica ciò significa che in certi SSD si possono ottenere valori per il VS di singolo membro inaccessibili per altri SSD (esempio SECS-P/07, 08, 09 contro FIS/04, M-PSI/02);*
- *La presenza di inattivi (-0.5 per ogni prodotto mancante) ha effetti “locali” che diventano sempre più importanti quanto più è alto il voto medio del SSD.*

## *Tuttavia:*

- *L'utilizzo della probabilità  $P_{inf}(U^*_d)$  come misuratore della qualità neutralizza in pratica questi effetti spuri residuali nel valore di  $U^*_d$ ;*
- *Su altre critiche apparse in rete, ho cercato di rispondere puntualmente, spero con garbo sufficiente (quasi sempre).*

# Una occhiata alle tabelle

*Le tabelle sono disponibili al sito ANVUR:*

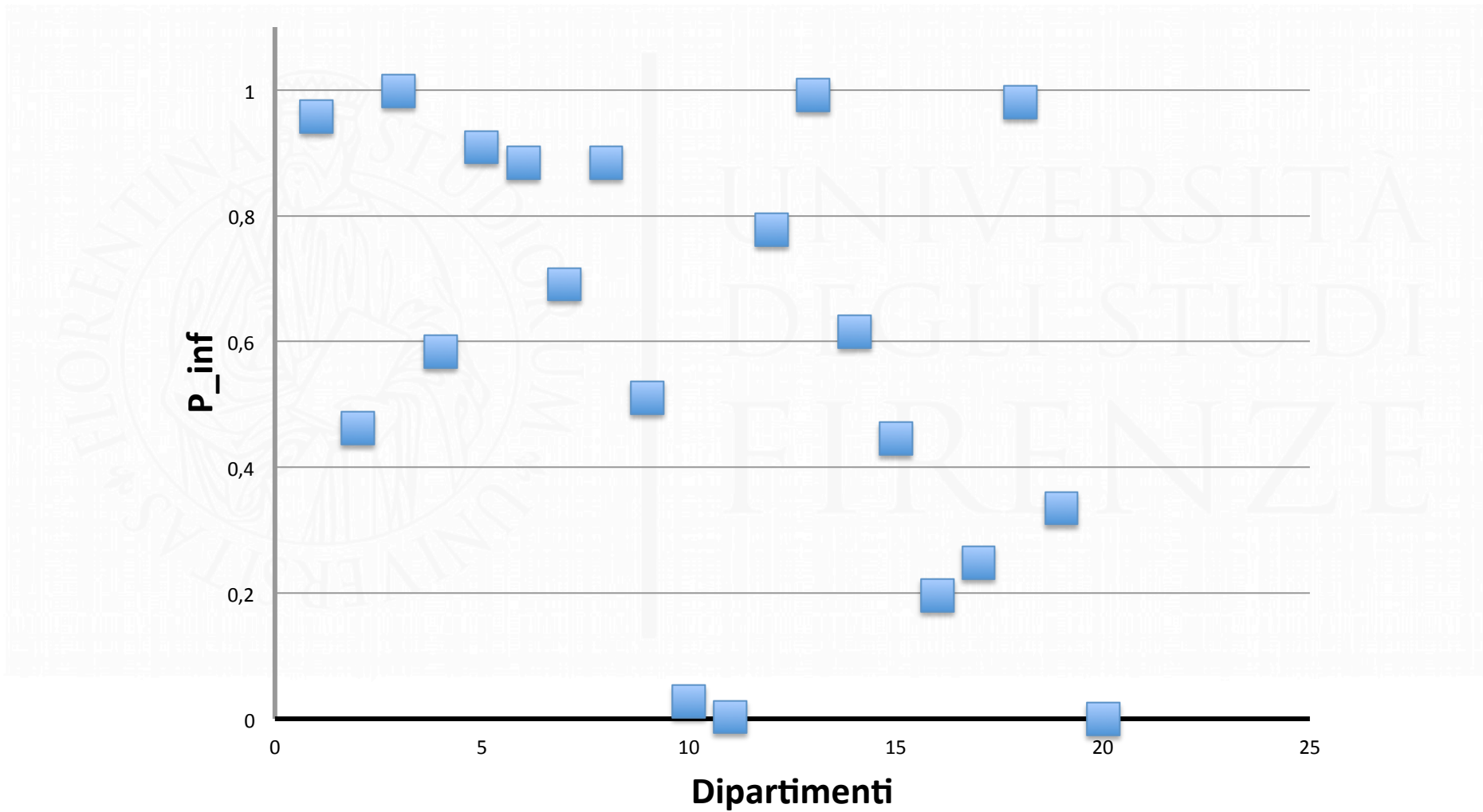
[http://www.anvur.org/index.php?option=com\\_content&view=article&id=609&Itemid=563&lang=it](http://www.anvur.org/index.php?option=com_content&view=article&id=609&Itemid=563&lang=it)

*Andamenti generali:*

- *distribuzione nazionale  $U_d^*$  e presenza di segnale*
- *Compensazione effetto dimensione*
- *Pesanti e preoccupanti differenze Nord /Centro /Sud /Isole*
- *Ogni dipartimento, almeno per i SSD con più di 10 prodotti, può vedere dalle tabelle quali sono i SSD che maggiormente hanno contribuito al risultato finale.*
- *Insistere sull'utilizzo dell'indicatore IPR derivato da VS per distribuire risorse non dipendenti dal costo delle ricerche*
- *Far notare i limiti nel caso di risorse di tipo differente*

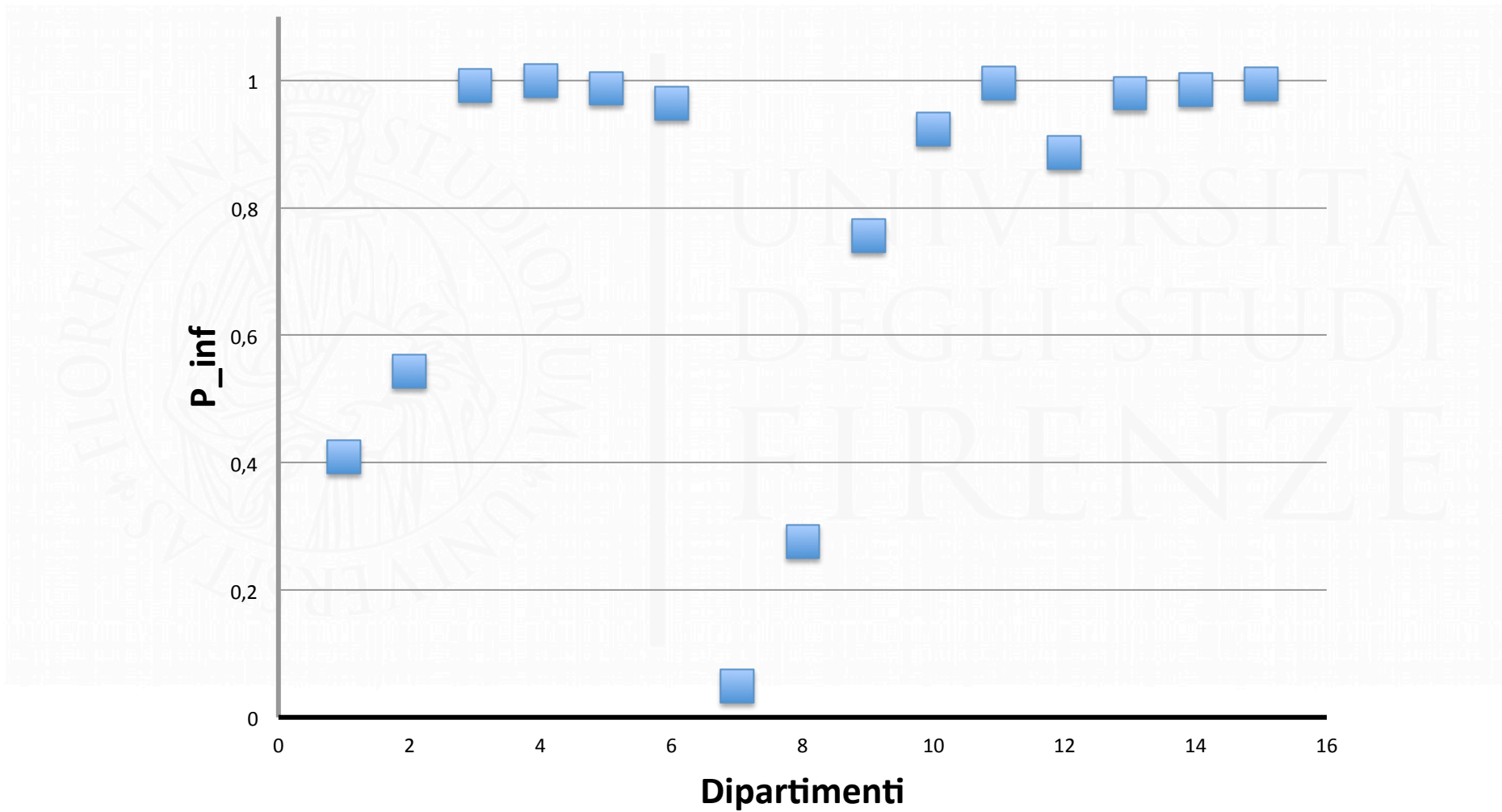
# UniPi

## Dipartimenti UniPi - P\_inf



# UniSi

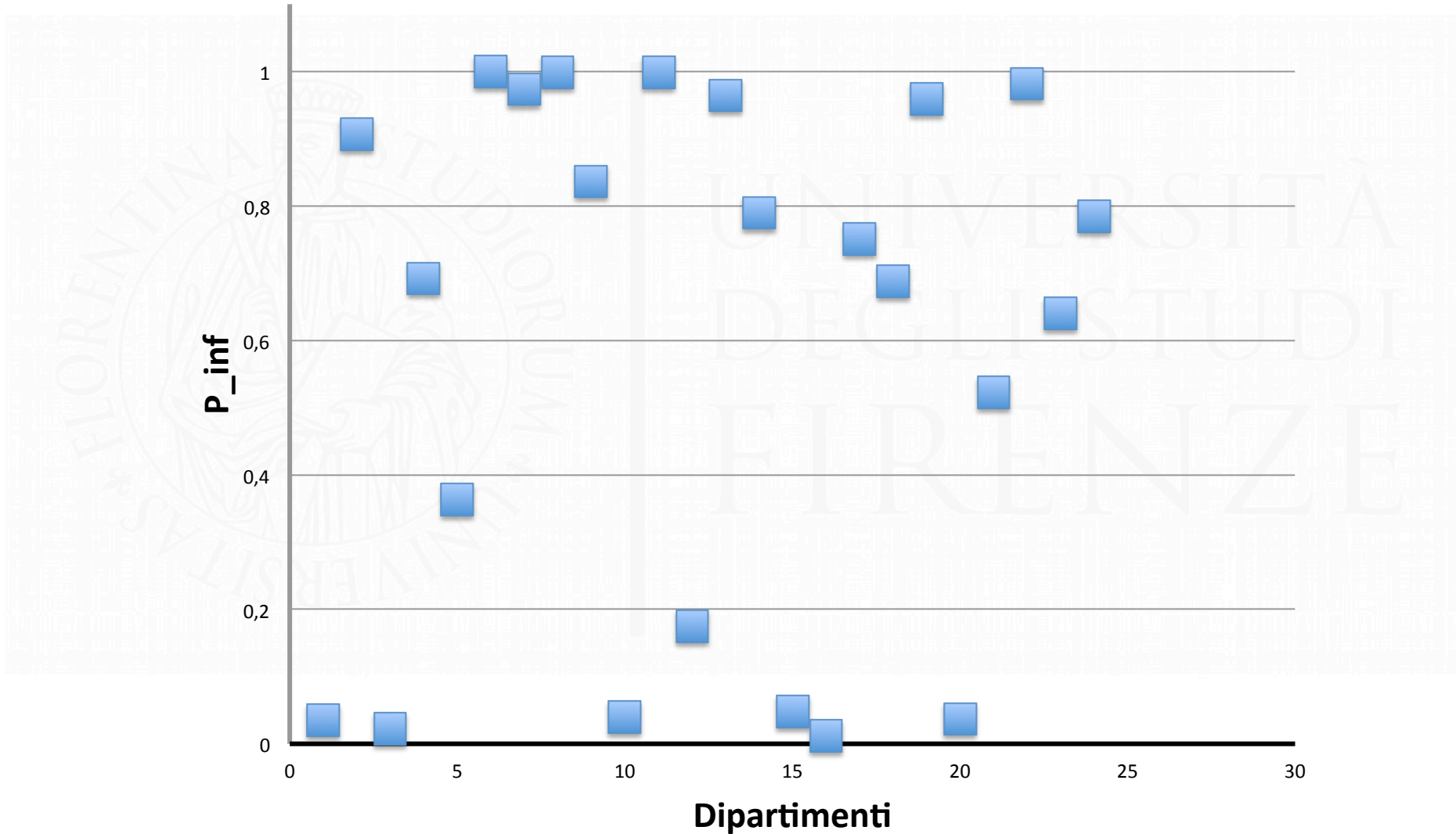
## Dipartimenti UniSi - P\_inf





# UniFi

## Dipartimenti UniFi - P\_inf



## Distribuire risorse non dipendenti dal costo ricerca

L'indicatore  $U_d$ , tramite  $P_{inf}(U_d^*)$ , definisce una relazione d'ordine quantitativa della qualità di ricerca dei dipartimenti.

La differenza  $IPR = [P_{inf}(U_d^*) - 0.5]$  può essere usata IPR per definire un indicatore quali-quantitativo  $US_d^*$  con cui distribuire risorse non dipendenti dal costo della ricerca:

$$IPR_d^* = P_{inf}(U_d^*) - 0.5$$

$$US_d^* = 1 + IPR_d^* \cdot \alpha$$

Il valore di  $\alpha$  è deciso dall'Ateneo, sulla base di decisioni "politiche".

Se l'Ateneo decide di distribuire le risorse con criteri prevalentemente premiali associati alla VQR, adotta valori di  $\alpha$  elevati; se invece predilige una distribuzione più a pioggia utilizza valori di  $\alpha$  piccoli.

## Distribuire risorse non dipendenti dal costo ricerca

*Le risorse totali RIS sono distribuite fra gli NDIP dipartimenti, assegnando al dipartimento d le risorse  $RIS_d$  :*

$$RIS_d = RIS \cdot \frac{US_d^* \cdot N_d}{\sum_{d=1}^{NDIP} US_d^* \cdot N_d}$$

*Potrebbe essere anche possibile distinguere la premialità (per dipartimenti con voto standardizzato positivo) dalle penalizzazioni (per dipartimenti con voto standardizzato negativo). Per far ciò si può usare l'indicatore  $USV_d$  definito come segue:*

$$USV_d^* = \begin{cases} 1 + IPR_d^* \cdot \alpha & \text{se } IPR_d^* \geq 0 \\ 1 + IPR_d^* \cdot \beta & \text{se } IPR_d^* < 0 \end{cases}$$

## Distribuire risorse non dipendenti dal costo ricerca

*Un Ateneo potrebbe decidere di non adottare alcun criterio premiale o limitativo delle risorse nei confronti di dipartimenti che l'Ateneo considera "normali" ovvero caratterizzati da valori di  $P_{inf}(U^*_d)$  che si collocano attorno a 0.5 e viceversa utilizzare l'indicatore IPR come detto sopra solo per i dipartimenti al di fuori di questo intervallo.*

*Necessario comunque che l'Ateneo richiami l'attenzione dei Dipartimenti a individuare al loro interno i contributi dei vari SSD. Per questo le tabelle che abbiamo fatto pubblicare dall'ANVUR contengono i VS di SSD (per quelli con più di 10 prodotti – clausola di riservatezza)*

# La collaborazione ANVUR-CRUI sugli indicatori VQR

*Azioni intraprese dal Gruppo di lavoro CRUI (ora coordinato da G. Ruocco):*

- *Chiesta all'ANVUR le valutazioni aggregate per SSD e non solo su Area CUN; l'ANVUR ha pubblicato tale informazione;*
- *Proposto all'ANVUR indicatori di qualità della ricerca dei Dipartimenti*
- *Il Voto Standardizzato è una di queste richieste*
- *L'ANVUR ha effettuato il calcolo (clausola di riservatezza)*
- *L'ANVUR pubblicherà a breve altre tabelle di VS di dipartimento, con altre condizioni (no soggetti inattivi, votazioni medie di tutti i prodotti presentati dal soggetto, inclusione dei pesi associati al costo della ricerca,...)*
- *Il Gruppo CRUI ha trovato estremamente positiva l'interazione con l'ANVUR su queste tematiche. Mi auguro che si tratti solo dell'inizio di una proficua e più ampia collaborazione.*

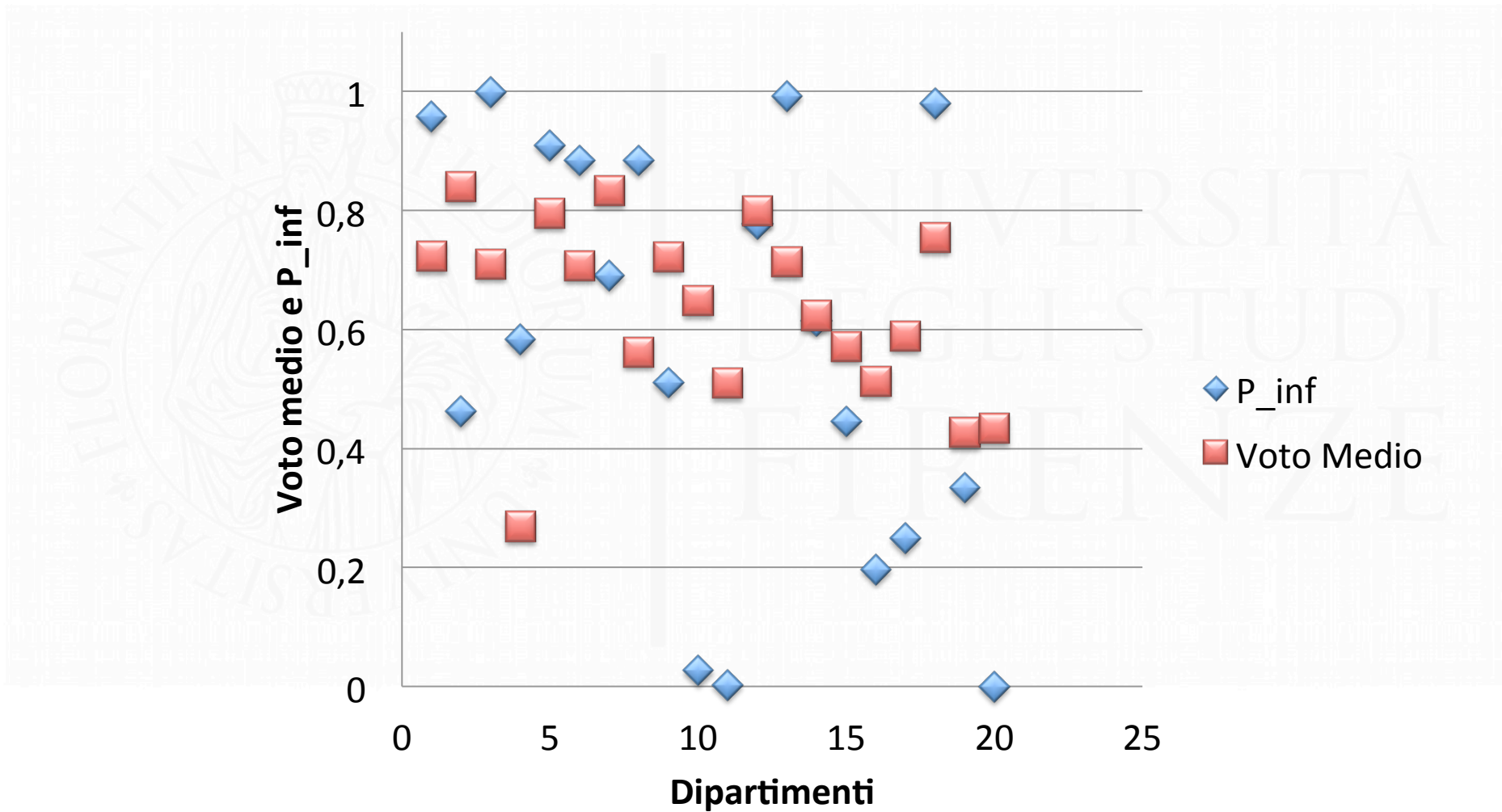
The seal of the University of Florence, featuring a seated figure holding a book, surrounded by the Latin text 'UNIVERSITAS FLORENTINA STUDIORUM' and 'UNIVERSITAS FLORENTINA STUDIORUM'.

*Grazie*

The logo of the University of Florence, consisting of the text 'UNIVERSITÀ DEGLI STUDI FIRENZE' in a serif font.

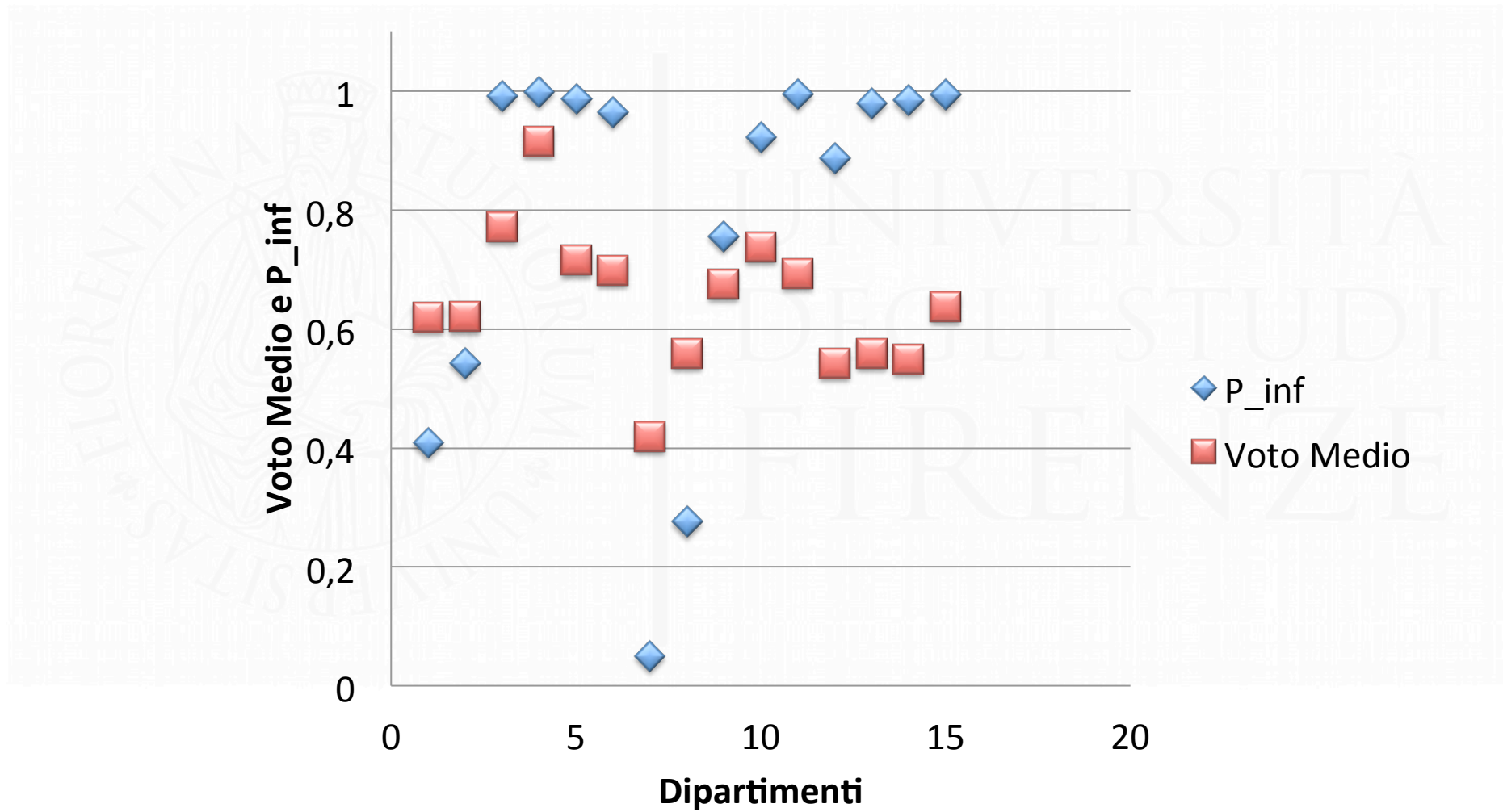
# UniPi

## Dipartimenti UniPi - Voto medio e P\_inf



# UniSi

## Dipartimenti UniSi - Voto Medio e P\_inf





# UniFi

## Dipartimenti UniFi - Voto Medio e P\_inf

