

Capitale umano

Introduzione

Quando si fa riferimento al termine “capitale”, inteso in un senso generico, spesso tendiamo ad accostarlo al solo e semplice capitale fisico. Impianti, macchine, computer e ogni altro mezzo facente parte del processo produttivo rientrano certamente in questo grande insieme, eppure da soli non bastano a delinearne i confini.

Accanto al capitale inteso in termini meramente fisici ce n'è infatti un altro, dai contorni forse (anzi certamente) più fumosi ma non per questo meno importante: si tratta del capitale umano.

All'interno di questa nuova definizione rientrano tutte quelle conoscenze, abilità e competenze individuali sia innate che acquisite dai lavoratori attraverso l'istruzione e la formazione professionale. Ma non solo, *“il capitale umano comprende tutto ciò che influenza la capacità degli individui di produrre e creare reddito, oltre alla forza delle loro braccia: la salute fisica e mentale ne è una determinante fondamentale”**.

Fermarsi solo all'istruzione appare quindi riduttivo, rischiando di omettere importanti componenti del processo di sviluppo della persona e del lavoratore quali possono essere ad esempio i fattori sanitari, fondamentali per interpretare in maniera corretta questa diversa accezione di capitale.

Oggetto di questo lavoro sarà dunque la relazione tra l'investimento in capitale umano e il tenore di vita nei diversi Paesi del mondo (misurato attraverso il PIL pro capite). Prendendo a riferimento i dati pubblicati dalla Banca Mondiale in tema di spesa per istruzione, formazione e sanità andremo quindi ad analizzare se ed in quale misura questo “investimento nell'individuo” possa portare ad un benessere dal punto di vista economico allo Stato che se ne fa promotore, il tutto con l'obiettivo costante di seguire un percorso il più possibile coerente e lineare.

** Intervento del Governatore della Banca d'Italia Ignazio Visco in occasione dei 15 anni di attività della Facoltà di Economia dell'Università Cattolica, sede di Roma*

Statistica descrittiva

Uno sguardo al Pil pro capite

Prima di iniziare l'analisi dell'influenza esercitata dall'investimento in capitale umano sul tenore di vita nei diversi Paesi è giusto focalizzarsi sulla variabile risposta di questa ricerca: il Pil pro capite.

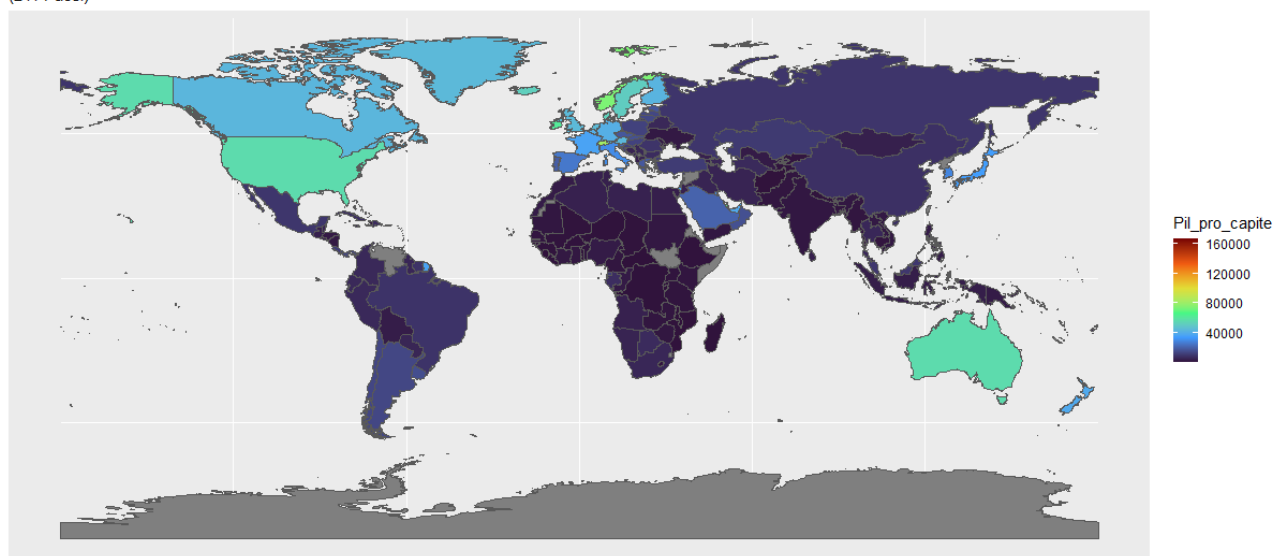
Come si distribuisce questa variabile a livello mondiale?

Si può osservare che la distribuzione del Pil non appare per nulla omogenea: infatti, ad una situazione caratterizzata da livelli di reddito pro capite molto bassi (soprattutto in Africa, Sud America e Asia), se ne contrappone un'altra speculare in cui il tenore di vita risulta essere decisamente più elevato (Europa, Nord America e Oceania).

Di particolare rilevanza sono i cosiddetti Microstati, ovvero quelle Nazioni aventi una popolazione ridotta e un territorio limitato, i quali occupano tutte le prime posizioni della classifica della distribuzione del Pil pro capite a livello mondiale. Difatti, il tenore di vita più elevato si ha nel Liechtenstein, seguito dal Principato di Monaco, dalle Bermuda, Lussemburgo e Isola di Man (tutte con un reddito pro capite maggiore di 85.000\$).

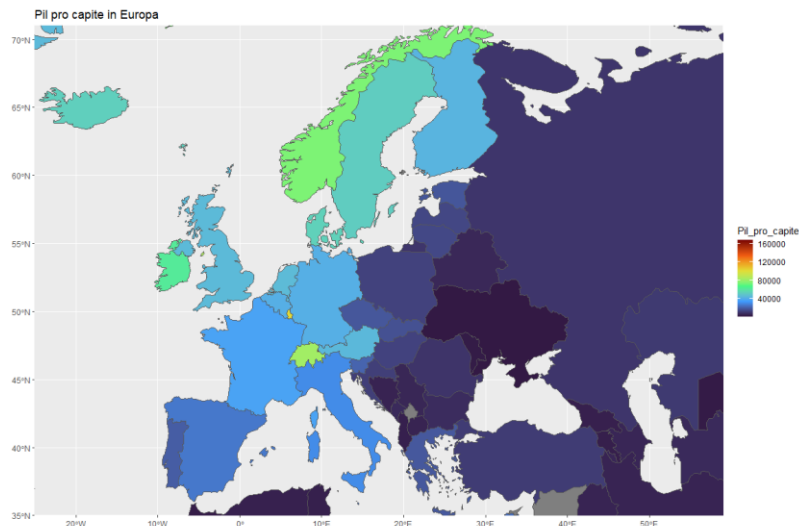
All'opposto, i Paesi dell'Africa centro-meridionale (Burundi, Rep. Centrale Africana, Malawi, Madagascar e Niger) risaltano all'occhio per i bassi standard quali-quantitativi del tenore di vita, con un livello di reddito pro capite addirittura al di sotto di 490\$.

Pil pro capite nel mondo
(241 Paesi)

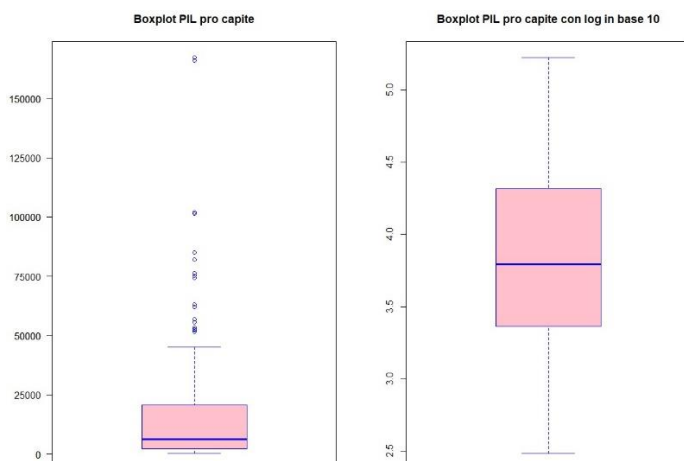
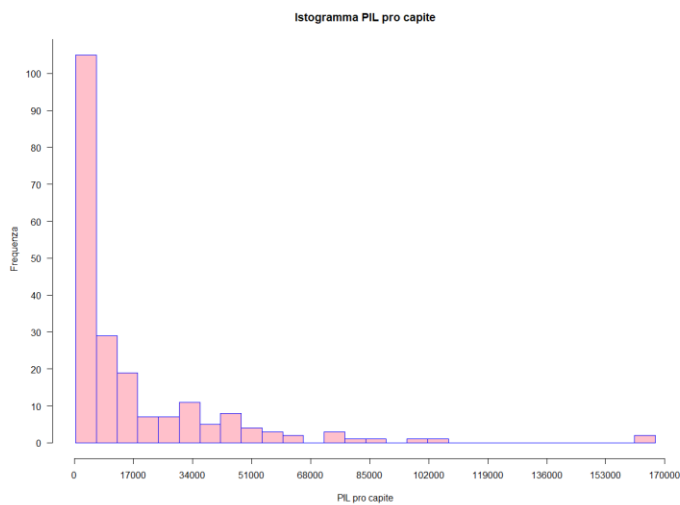


A livello europeo il tenore di vita risulta invece essere decisamente più omogeneo che a livello mondiale, sebbene sia comunque presente una differenza tra i Paesi del Nord e del Sud Europa. In particolare si distinguono per gli elevati livelli di Pil pro capite i Paesi scandinavi (insieme alla Svizzera, Irlanda e ai Microstati), seguiti dai Paesi dell'Europa continentale e Meridionale, mentre i Paesi dell'area Balcanica rimangono ancorati a bassi livelli di reddito.

L'Italia, con un Pil pro capite di circa 30.200\$, si rivela essere uno degli Stati con gli standard più elevati nel Sud Europa, sebbene permangano delle leggere differenze con le Nazioni dell'area Continentale quali Francia (36.638\$), Germania (41.086\$), Austria (44.178\$) e Regno Unito (44.974\$).



I grafici riportati di seguito mostrano come la distribuzione del reddito per abitante non abbia un profilo per nulla simmetrico e, anzi, presenti un'evidente asimmetria positiva. Il 50% dei Paesi si colloca infatti tra un livello di Pil pro capite compreso tra i 2.132,2\$ e i 17.879,5\$, rendendo così le frequenze delle classi di reddito più basse sia quantitativamente che visivamente più rilevanti delle classi più alte.

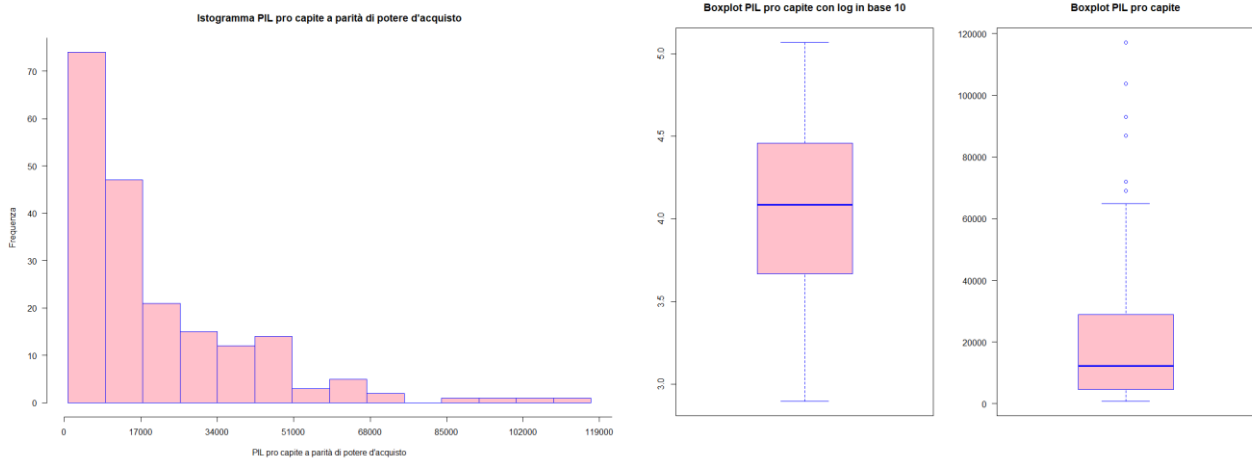


Il boxplot della distribuzione a livello mondiale del Pil pro capite: si evidenzia come la maggior parte dei Paesi si concentri in fasce basse di reddito per abitante, mentre spostandosi verso l'alto sia presente un elevato numero di outliers dato soprattutto dai Microstati. A fianco un boxplot realizzato utilizzando il logaritmo in base 10 dei livelli di Pil pro capite.

Focus: il Pil pro capite a parità di potere d'acquisto

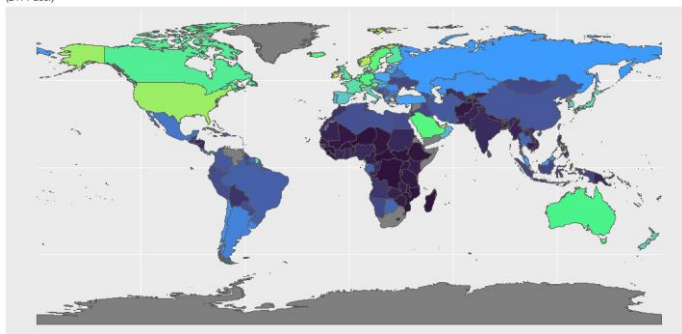
Presentato il modo di distribuirsi del Pil pro capite tra i vari Paesi del mondo, si anticipa che nel corso della trattazione verrà utilizzata una versione particolare di questa variabile: il Pil pro capite a parità di potere d'acquisto. Questa misura tiene infatti conto delle differenze nel livello generale dei prezzi tra i diversi Stati.

Anche per questa variabile, per completezza, vengono presentati i grafici già analizzati nello studio del Pil pro capite.



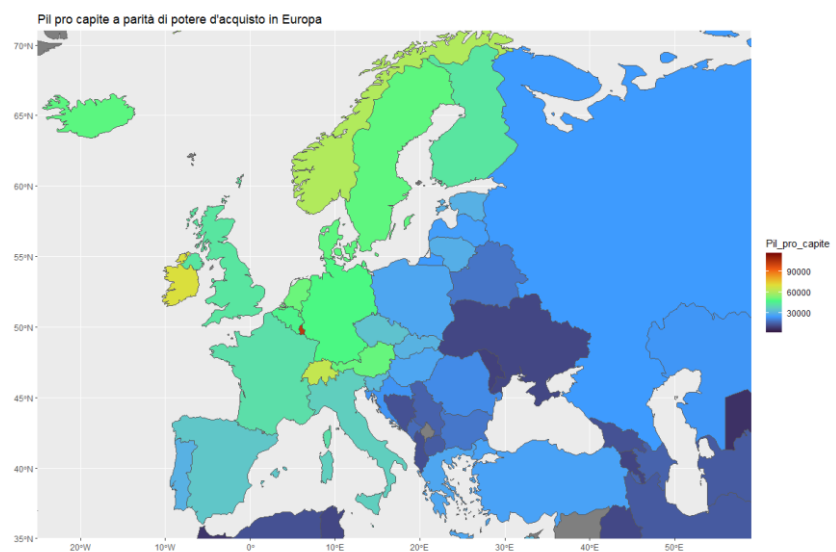
Istogramma e boxplot (anche in logaritmi in base 10) del Pil pro capite PPA

Pil pro capite a parità di potere d'acquisto nel mondo
(241 Paesi)



Distribuzione a livello mondiale del Pil pro capite a parità di potere d'acquisto

Distribuzione a livello europeo del Pil pro capite a parità di potere d'acquisto



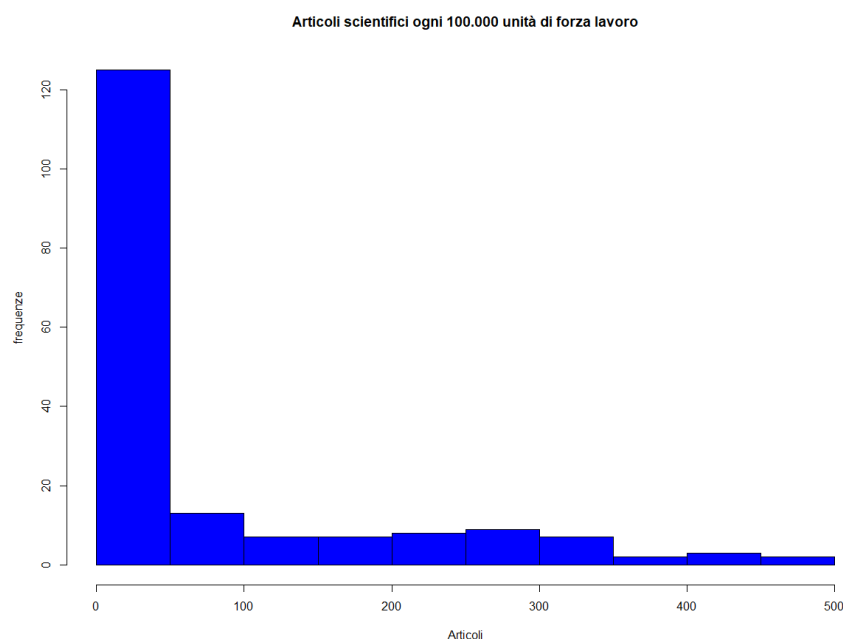
Due misure dell'investimento in capitale umano

Si intendono ora presentare due variabili rappresentative del livello di investimento in capitale umano nei diversi Paesi del mondo: *gli articoli scientifici ogni 100.000 unità di forza lavoro* e il *rapporto alunni-docenti nella scuola primaria*. Poiché queste misure ricorreranno nel corso della trattazione, verranno di seguito sinteticamente analizzate le loro principali caratteristiche e la loro distribuzione a livello mondiale.

Gli articoli di riviste scientifiche e tecniche

La pubblicazione scientifica è una misura importante nel tentativo di spiegare il differente contributo di ciascun Paese allo sviluppo del settore della ricerca e all'innovazione tecnologica, ma non solo, essa riflette anche l'impegno dei diversi Stati volto all'incremento del pool di conoscenze scientifiche e di competenze tecniche all'interno dei propri confini. Per "articoli di riviste scientifiche e tecniche" si intende così il numero di articoli scientifici e ingegneristici pubblicati nei seguenti campi: fisica, biologia, chimica, matematica, medicina clinica, ricerca biomedica, ingegneria e tecnologia e scienze della terra e dello spazio.

Poiché tale misura privilegiava però in una qualche maniera gli Stati più popolosi a danno di quelli con una popolazione ridotta, si è deciso di rapportare tale variabile alla forza lavoro presente nei diversi Paesi e successivamente di moltiplicarla per 100.000. Mentre infatti la grandezza originaria assumeva i suoi valori più elevanti in corrispondenza di Usa (circa 429.988 articoli), Cina (circa 407.974 articoli), Germania, India e Giappone, questa nuova misura tiene conto delle differenze riguardanti l'entità della popolazione nel mondo del lavoro nei diversi contesti di analisi, presentando così valori molto elevati per Paesi come la Danimarca (circa 492 articoli ogni 100.000 unità di forza lavoro), la Svizzera, l'Australia, la Finlandia e la Svezia. Inoltre, lì dove la variabile originaria presentava i suoi valori più bassi in presenza di Stati insulari con una popolazione molto ridotta (ma non per questo con un ridotto impegno nella ricerca scientifica), gli articoli scientifici ogni 100.000 unità di forza lavoro assumono i loro valori minori in corrispondenza del Sudan del Sud, della Somalia, del Chad, della Nord Corea e dell'Afghanistan, che contribuiscono proporzionalmente di meno allo sviluppo e all'innovazione tecnologica.



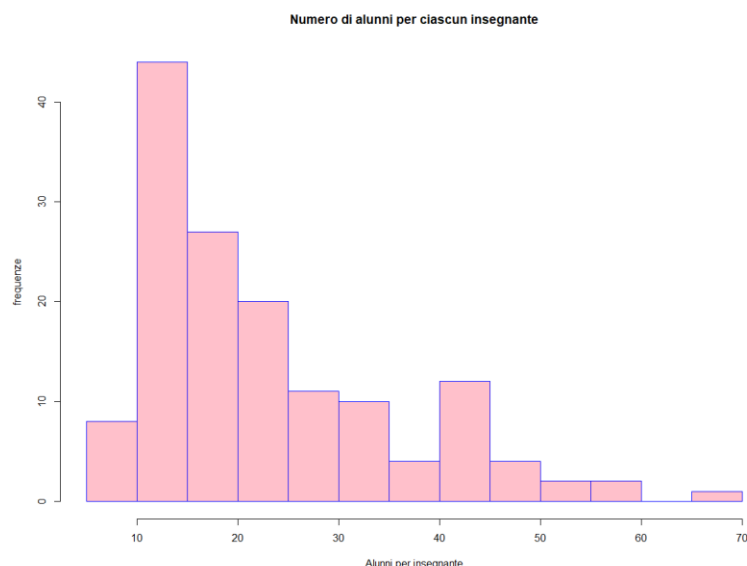
Si evidenzia anche qui, così come in precedenza nell'analisi della distribuzione del Pil pro capite, una evidente asimmetria positiva.

Difatti su un gruppo di 220 Paesi più della metà si trova nella classe 0-50, evidenziando così una diffomità significativa, sebbene intuibile, nell'apporto alla ricerca e allo sviluppo tecnologico dei differenti Paesi a livello mondiale.

Istogramma di frequenza del numero di articoli scientifici ogni 100.000 unità di forza lavoro

Il rapporto alunni-docenti nella scuola primaria

L'ultima grandezza che si intende presentare prima di iniziare la trattazione riguardo l'inferenza statistica è il rapporto alunni-docenti nella scuola primaria, ovvero il numero medio di alunni per singolo insegnante nei vari Paesi del mondo. Avendo presentato in precedenza una misura del contributo e dell'investimento dei diversi Stati nel settore della ricerca si tenta ora, mediante tale rapporto, di dare un'immagine della qualità dei differenti sistemi scolastici. Il focus sarà così indirizzato verso una particolare tipologia di investimento in capitale umano: quello verso i primi gradi di istruzione.

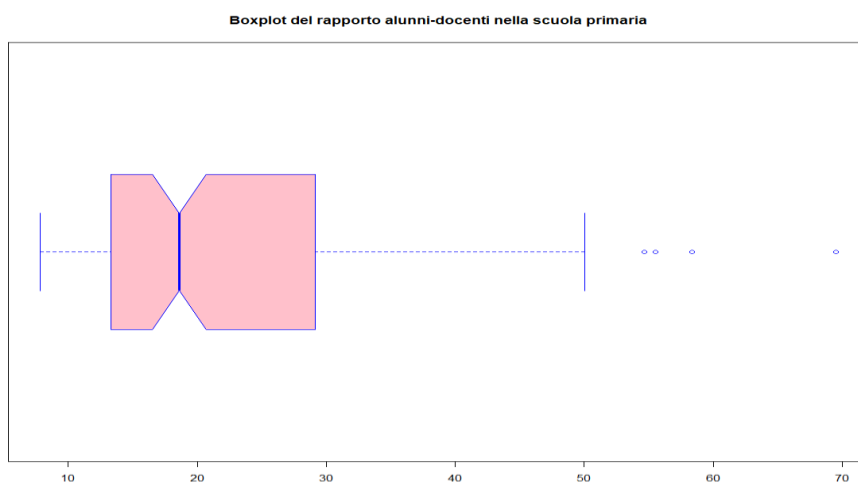


Anche tale distribuzione presenta, come le due precedenti, un'asimmetria positiva, avendo la coda tendente verso destra. L'interpretazione che se ne dà è tuttavia differente: mentre in precedenza bassi livelli di Pil pro capite o di pubblicazioni scientifiche segnalavano un certo livello di inefficienza, qui un ridotto rapporto alunni-insegnanti è invece sintomo della qualità del sistema scolastico a livello primario dei singoli Paesi.

Istogramma di frequenza del numero di alunni per insegnante nella scuola primaria

Le Nazioni con un più elevato rapporto alunni-docenti risultano così essere ancora una volta quelle del continente africano, tra cui il Malawi (con una media di 69,5 alunni per insegnante), il Rwanda (con una media di circa 58,3), il Chad, il Mozambico o l'Angola. All'opposto il rapporto tende ad assumere i suoi valori minimi in presenza soprattutto di Microstati come il Liechtenstein e il Lussemburgo (entrambi con una media minore di 8,5 alunni per docente) o dei Paesi del Nord Europa tra cui ad esempio la Norvegia (con circa 8,9 alunni per docente).

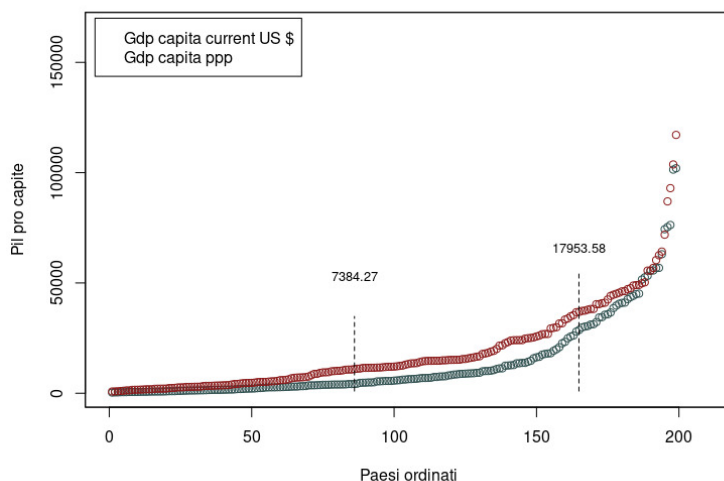
Il boxplot riportato qui di fianco evidenzia anch'esso l'asimmetria positiva già evidenziabile dall'istogramma. Il 50% dei dati è compreso nell'intervallo 13,3-29,1.



Statistica inferenziale

Pil pro capite e Pil pro capite a parità di potere d'acquisto a confronto

Come anticipato nella parte riguardante la statistica descrittiva, nel corso della trattazione si utilizzerà una misura particolare del Pil pro capite, ossia quella che tiene conto del livello generale dei prezzi nei vari Paesi del mondo. Si intende così introdurre questa nuova fase del lavoro mediante un breve confronto tra il Pil pro capite e il Pil pro capite a parità di potere d'acquisto.



Il grafico a sinistra mette così in evidenza l'andamento delle due diverse variabili a partire dai Paesi più poveri sino a quelli tendenzialmente più ricchi, identificando con il colore verde il *Pil pro capite* e con il colore rosso il *Pil pro capite PPP*. L'andamento delle due misure risulta essere il medesimo, sebbene il Pil pro capite a parità di potere d'acquisto assuma valori leggermente più elevati soprattutto in corrispondenza di Stati con un livello medio di reddito per abitante.

Un'ultima assunzione rilevabile dal grafico: si è deciso dividere le varie Nazioni oggetto dello studio in tre categorie, ossia in Paesi poveri (fino a 7.384,27\$ di Pil pro capite PPP), Paesi intermedi (7.384,27\$-17.953,58\$) e Paesi ricchi (più di 17.953,58\$): nel seguito verranno così utilizzati i Paesi poveri nell'ambito della stima bayesiana e i Paesi medio-ricchi nel modello di regressione.

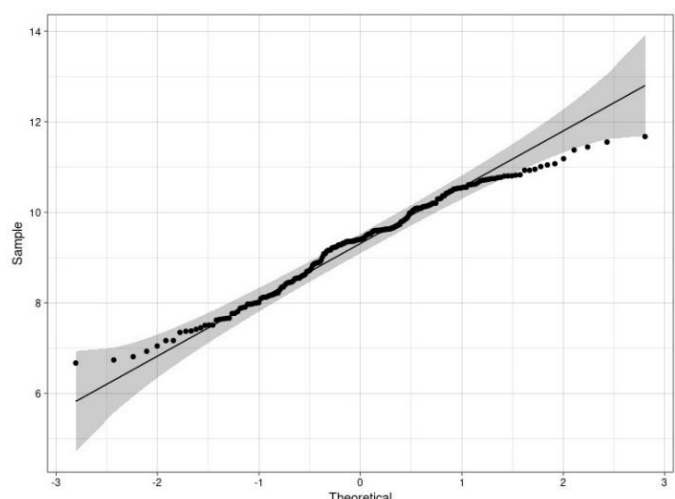
La distribuzione del Pil pro capite PPP: test di Shapiro-Wilk e test di Kolmogorov-Smirnov

Volendo analizzare quale potesse essere la forma della variabile casuale generatrice dei dati campionari in nostro possesso riguardanti il Pil pro capite PPP si è inizialmente voluta testare l'ipotesi di provenienza da una popolazione normale tramite il test di Shapiro-Wilk ed una contestuale verifica grafica tramite il qqplot. Pur avendo espresso la variabile di nostro interesse in logaritmi (tentando così di ridurre le differenze tra Paesi ricchi e poveri), il test non ha dato esito positivo.

Shapiro-Wilk normality test

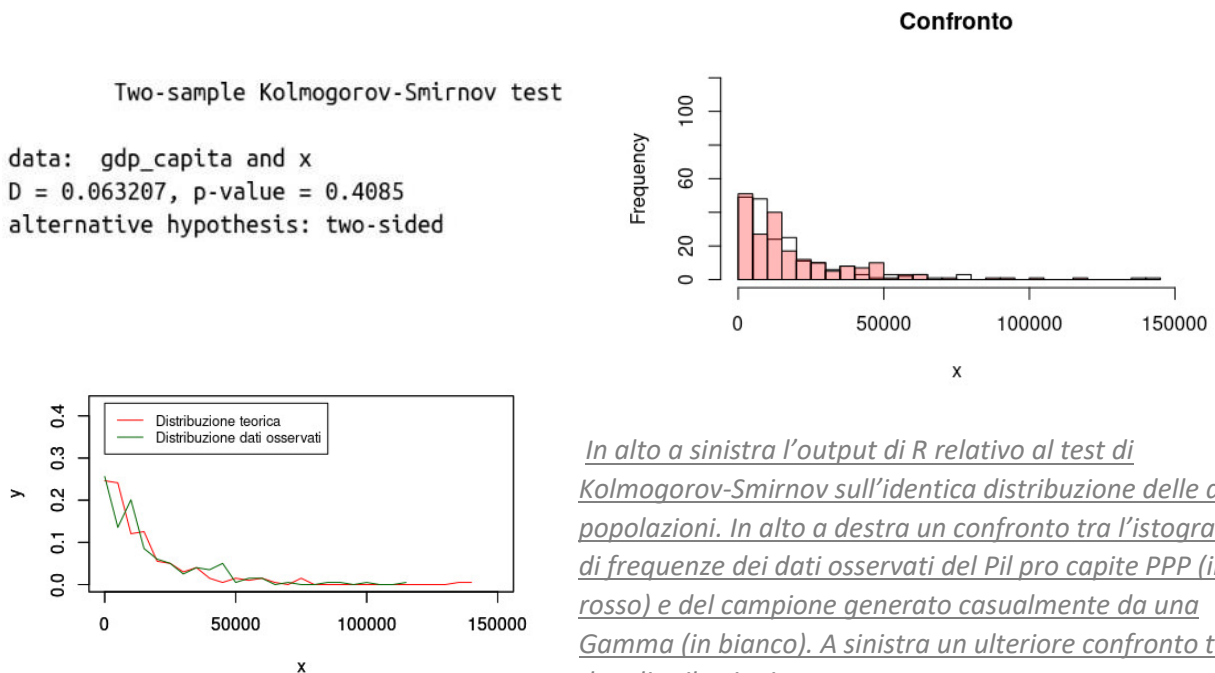
```
data: log(gdp_capita)
W = 0.97946, p-value = 0.005142
```

Output di R del test di Shapiro-Wilk: il p-value risulta minore di 0,05 per cui si rifiuta l'ipotesi nulla di distribuzione normale della popolazione. Di fianco il qq-plot, nel quale si può osservare come i punti divergano dalla diagonale, specialmente nelle due code.

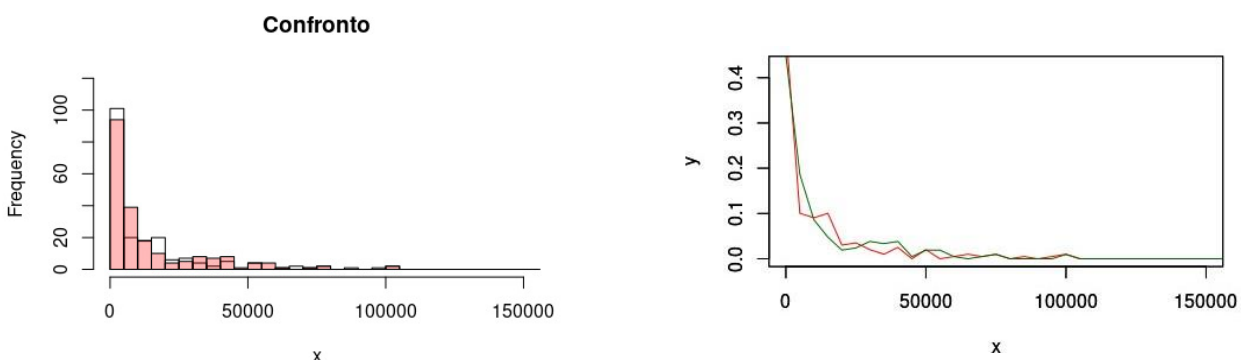


Rifiutata l'ipotesi di provenienza da una popolazione normale e ricordata l'asimmetria positiva nella distribuzione del reddito pro capite PPP osservata precedentemente, l'analisi si è così focalizzata sulla verifica della provenienza del campione da una variabile casuale Gamma. Per poter testare questa nuova ipotesi si è però reso necessario calcolare dapprima i parametri della variabile Gamma che si è supposto poter essere la generatrice dei nostri dati. Sono così stati calcolati i parametri shape e scale, i quali risultano essere rispettivamente *il rapporto tra media al quadrato e varianza del Pil pro capite PPP* (pari a 0,9307997) e *il rapporto tra varianza e media della medesima variabile* (pari a 20.355,12).

Per la verifica della provenienza del campione da una variabile casuale Gamma è stato inizialmente utilizzato il test di adattamento di Kolmogorov-Smirnov, il quale tuttavia non ha dato un esito positivo, sebbene le due distribuzioni sembrassero in prima analisi molto simili. Poiché il problema pareva essere dato più che altro dal basso numero di osservazioni in nostro possesso, dopo aver generato un campione casuale dalla distribuzione Gamma con i parametri stimati in precedenza, si è proceduto così ad un nuovo test di Kolmogorov-Smirnov per testare l'identica distribuzione delle popolazioni generatrici dei due campioni: tale test ha dato esito positivo. Per sicurezza il test è stato ripetuto confrontando i dati in nostro possesso con campioni generati casualmente da altre variabili casuali, ma in tutti questi casi i bassi livelli del p-value hanno portato a rifiutare l'ipotesi di identica distribuzione delle popolazioni.



Si osserva inoltre che anche il confronto tra la variabile "Pil pro capite" (non PPP) ed un campione casuale generato da una Gamma (con parametri shape e scale stimati utilizzando la medesima formula) presenta una similitudine tra le due distribuzioni.



Stima puntuale

Dopo aver analizzato la distribuzione della popolazione di riferimento dei dati in nostro possesso, il prossimo tema ad essere presentato sarà quello della stima puntuale. Verranno così introdotti separatamente gli stimatori BLUE di media e varianza e gli stimatori MLE dei parametri shape e scale per i Paesi medio-ricchi e per i Paesi poveri, presentando poi per questi ultimi un approccio di stima puntuale di tipo bayesiano.

Gli stimatori BLUE di media e varianza per i Paesi medio-ricchi

Si presentano inizialmente gli stimatori BLUE di media e varianza di quella fascia di Paesi in precedenza definiti medio-ricchi (Pil pro capite PPP maggiore di 7.384,27\$). Poiché tali stimatori presumono unicamente la conoscenza del campione, non è stata fatta alcuna congettura riguardo la distribuzione della popolazione da cui provengono i nostri dati.

La media e la varianza campionaria del Pil pro capite PPP in tale caso risultano essere rispettivamente di 32.280\$ e di 326.000.000 (dollari al quadrato).

Di fianco viene invece riportato l'output di R contenente gli stimatori BLUE della media e della varianza nella popolazione; essendo inoltre tali stimatori non distorti, ne vengono anche presentate le varianze (terza riga e quarta riga) che in questo caso coincidono con gli errori quadratici medi.

```
> (BLUE_mean <- BLUE_est$muhat)
      [,1]
[1,] 31278.01
> (BLUE_sigma <- BLUE_est$sigmahat^2)
      [,1]
[1,] 289320381
> (BLUE_var_mu <- sqrt(BLUE_est$Var_mu))
      [,1]
[1,] 2919.386
> (BLUE_var_var <- sqrt(BLUE_est$Var_sigma))
      [,1]
[1,] 3104.84
```

Gli stimatori BLUE di media e varianza per i Paesi poveri

```
> (BLUE_mean.poor <- BLUE_est.poor$muhat)
      [,1]
[1,] 3848
> (BLUE_sigma.poor <- BLUE_est.poor$sigmahat)
      [,1]
[1,] 1262
> (BLUE_var_mu.poor <- BLUE_est.poor$Var_mu)
      [,1]
[1,] 54505
> (BLUE_var_var.poor <- BLUE_est.poor$Var_sigma)
      [,1]
[1,] 37878
```

A sinistra l'output di R riguardante gli stimatori BLUE di media e varianza per la fascia di Paesi poveri. Nella terza e quarta riga sono presentati, come in precedenza, gli errori quadratici medi.

Gli stimatori MLE dei parametri shape e scale per i Paesi medio-ricchi

Per la stima dei parametri shape e scale è stato invece utilizzato il *metodo della massima verosimiglianza (MLE)*, in cui, oltre al campione, si presume anche la conoscenza della distribuzione della popolazione generatrice delle osservazioni campionarie.

```
mle_est_rich <- fitdistr(dati, "gamma", list(shape=3.19, scale=10105), lower=0.01)
mle_est_rich
  shape      scale
3.24e+00  1.01e+04
(7.57e-01) (2.58e+03)
```

Ipotizzando dunque, in base ai precedenti risultati raggiunti con il test di Kolmogorov-Smirnov, che i dati campionari siano generati da una variabile Gamma si è proceduto alla stima MLE dei parametri, la quale, mediante il comando *fitdistr()*, richiede le osservazioni campionarie (*dati*), la popolazione generatrice dei

dati campionari (*Gamma*) e un valore iniziale dei parametri da stimare. Al di sotto degli stimatori MLE sono invece presentate le variazioni della stima.

Gli stimatori MLE dei parametri shape e scale per i Paesi medio-ricchi

Anche per i Paesi poveri, avendo seguito il medesimo procedimento appena presentato, vengono riportati gli stimatori MLE dei parametri shape e scale. Come per i Paesi ricchi si è assunta la provenienza dei dati campionari da una popolazione distribuita come una variabile casuale Gamma.

```
mle_est <- fitdistr(c(data), "gamma", list(shape=3.95, scale=990), lower = 0.1)
mle_est
  shape    scale
  3.2    1217.8
  1.6) ( 671.3)
```