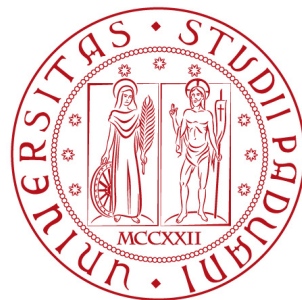


Università degli Studi di Padova  
Dipartimento di Scienze Statistiche  
Corso di Laurea Triennale in  
Statistica per l'Economia e l'Impresa



RELAZIONE FINALE

**ANALISI STATISTICA DELLA VALUTAZIONE  
DEI NERVI CRANICI E DELLA DISFAGIA  
IN PAZIENTI AFFETTI DA SLA**

**Relatore:** Prof.ssa Laura Ventura

**Correlatore:** Dott.ssa Irene Battel

**Laureando:** Silvia Montaldo

**Matricola:** 1124655

Anno Accademico 2017/2018



# Indice

<b>1</b>	<b>INTRODUZIONE</b>	<b>1</b>
1.1	La Sclerosi Laterale Amiotrofica . . . . .	1
1.2	La disfagia nella SLA . . . . .	3
1.3	Obiettivi dello studio . . . . .	5
1.4	I dati . . . . .	6
1.5	Conclusioni . . . . .	9
<b>2</b>	<b>ANALISI ESPLORATIVE PRELIMINARI</b>	<b>11</b>
2.1	Analisi univariate . . . . .	11
2.2	Analisi bivariate . . . . .	17
2.3	Conclusioni . . . . .	29
<b>3</b>	<b>LA CURVA ROC</b>	<b>31</b>
3.1	Introduzione . . . . .	31
3.2	Indici sintetici . . . . .	32
3.3	La curva ROC . . . . .	34
3.4	Applicazione della curva ROC al dataset . . . . .	36
3.5	Conclusioni . . . . .	41
<b>4</b>	<b>MODELLI DI REGRESSIONE LOGISTICA</b>	<b>43</b>
4.1	La regressione logistica . . . . .	43
4.2	Il problema di separazione e la correzione di Firth . . . . .	45
4.3	Applicazione del metodo di Firth al dataset . . . . .	47
4.4	Conclusioni . . . . .	51



# Capitolo 1

## INTRODUZIONE

### 1.1 La Sclerosi Laterale Amiotrofica

La Sclerosi Laterale Amiotrofica (SLA) è una malattia degenerativa del sistema nervoso centrale caratterizzata dal coinvolgimento contemporaneo del primo motoneurone (quello corticale o centrale) e del secondo motoneurone (quello periferico, posto a livello delle corna anteriori del midollo spinale o dei nuclei motori degli ultimi nervi cranici).

La SLA fu descritta fra il 1869 e il 1874 da Jean-Martin Charcot, neurologo francese; nel 1932 Brain introdusse il termine di malattia del motoneurone (motor neuron disease, MND), per indicare un insieme di patologie aventi quale comune denominatore la compromissione primaria dei neuroni di moto centrali e/o periferici (Mutani *et al.*, 2007).

Nei paesi occidentali, l'incidenza della SLA è compresa fra 1.7 e 2.6 casi per 100.000 abitanti all'anno; la prevalenza è pari a circa 6-8 casi per 100.000 abitanti; la frequenza sembra essere in lieve aumento, in parte a causa dell'incremento dell'età media della popolazione. La SLA è lievemente più comune nel sesso maschile (rapporto maschi/femmine pari a 1.3). Si tratta di una malattia tipica dell'età adulta, difatti l'età media di esordio è di circa 65 anni, tuttavia sono relativamente frequenti casi con esordio in età giovanile (sotto i 45 anni), che presentano un decorso clinico assai più benigno rispetto ai casi con esordio in età adulta.

Tra i fattori di rischio noti o sospetti si ha il fumo di sigaretta, l'attività agricola e i traumi, ma la causa di tale malattia è tuttora sconosciuta (Mutani *et al.*, 2007).

La caratteristica patologica fondamentale della SLA è la rapida e progressiva perdita dei neuroni motori superiori ed inferiori: la degenerazione di tali neuroni motori provoca una paralisi progressiva che arriva a coinvolgere i muscoli respiratori, portando alla morte o al ricorso a ventilazione meccanica.

Le manifestazioni cliniche della SLA sono il risultato della compromissione del motoneurone centrale e/o periferico, ma i segni e i sintomi della malattia, in particolare all'esordio, possono essere così eterogenei da rendere spesso difficoltosa una diagnosi di certezza.

L'esordio può essere di tipo spinale o di tipo bulbare anche se in molti casi, specie nella fase iniziale, si presenta con sintomi aspecifici.

Nella forma spinale vengono colpiti per primi, o in prevalenza, i motoneuroni spinali, ossia quelli localizzati a livello del midollo spinale. La maggior parte dei pazienti (circa il 70%) presenta un esordio di questo tipo, caratterizzato frequentemente da debolezza muscolare e atrofia focale degli arti inferiori e superiori. Nella fase avanzata della malattia questi sintomi tendono a progredire fino alla paralisi degli arti e a ciò si associano anche, nella maggior parte dei pazienti, sintomi di tipo bulbare.

Il restante 30% dei pazienti, presenta un esordio di tipo bulbare; in questa forma vengono colpiti per primi, o in prevalenza, i motoneuroni inferiori localizzati nella regione bulbare del tronco encefalico (Wijesekera e Leigh, 2009). La SLA ha un decorso progressivamente ingravescente e di solito porta a morte in 2-3 anni per complicazioni varie o per insufficienza respiratoria da interessamento del diaframma, vi sono tuttavia casi a durata molto più lunga, fino a 15-20 anni.

Tra i fattori prognostici negativi (indice di un decorso rapido) troviamo l'insorgenza in età avanzata o con sintomi bulbari e la presenza di un deficit respiratorio precoce. Non esiste attualmente alcuna terapia in grado di arrestare la progressione della malattia né di farne regredire sintomi e segni. Tuttavia, anche se la SLA non è una malattia guaribile, sono disponibili mol-

ti interventi in grado di alleviare le sofferenze del paziente e preservarne la qualità di vita.

Questo studio si concentra sui problemi legati alla nutrizione del paziente affetto da SLA, il quale può presentare frequentemente turbe della deglutizione e alterazioni dello stato nutrizionale, causate dalla riduzione delle ingesta per os (cioè per via orale).

La riduzione dell'alimentazione per os avviene progressivamente: il paziente tende a ridurre la sua alimentazione gradualmente, in seguito al progressivo aggravarsi delle condizioni funzionali e ai disagi soggettivi conseguiti alla necessità di alimentarsi con cibi di consistenza differente dall'usuale e/o all'aumento del tempo e della fatica necessari per deglutire il pasto.

Se l'alimentazione orale non è sufficiente a coprire le richieste energetico-proteiche del paziente, si può associare l'uso di integratori alimentari presenti in commercio o iniziare la nutrizione enterale, che rappresenta il sistema nutrizionale di scelta quando è presente disfagia grave e sia necessario il ricorso all'alimentazione artificiale. La nutrizione enterale può avvenire mediante sondino naso-gastrico, metodica oggi per lo più abbandonata nella SLA, o mediante sondino gastrico inserito per via gastroscopica (PEG) o radiologica (RIG o PRG). In pazienti con grave compromissione respiratoria in cui l'inserimento di un sondino gastrico sia controindicato ci si può avvalere della nutrizione parenterale, nella quale la soluzione nutritiva viene somministrata per infusione attraverso una linea venosa centrale (Mutani *et al.*, 2007).

Da un punto di vista clinico, questa distinzione è utile specialmente nelle fasi iniziali della malattia e non durante l'evoluzione della stessa, in quanto, nella maggior parte dei casi, le due forme tendono a sovrapporsi (Rozzani, 2017).

## 1.2 La disfagia nella SLA

La deglutizione è un atto neuromotorio composto da una serie di movimenti che permettono il trasporto del cibo dalla cavità orale allo stomaco, evitando il passaggio di materiale alimentare nelle vie aeree.

Una deglutizione normale richiede la precisa coordinazione di circa 30 muscoli localizzati nella cavità orale, nella faringe, nella laringe e nell'esofago. Tali muscoli sono controllati da 5 paia di nervi cranici (V Trigemino; VII Faciale; IX Glossofaringeo; X Vago; XII Ipoglosso) e da 3 paia di nervi spinali cervicali (C1 – C3) e la loro coordinazione è determinata dal centro della deglutizione, localizzato nel tronco encefalico a livello della medulla allungata (Dodds *et al.*, 1990).

Con il termine disfagia si intende un'alterazione del meccanismo di deglutizione che può essere di tipo neurogeno, quando è associata a lesioni neurologiche, o di tipo meccanico, quando è associata ad alterazioni anatomiche.

La disfagia neurogena può essere definita come la difficoltà a deglutire a causa di una patologia neurologica, che interferisce con il transito del cibo dalla cavità orale allo stomaco. La sintomatologia è molto varia: vengono normalmente distinti dei sintomi diretti di disfagia (tosse, necessità di effettuare deglutizioni multiple, rallentamento che coinvolge le diverse fasi, voce umida o gorgogliante) e dei sintomi indiretti (calo ponderale, disidratazione, febbri ricorrenti, frequenti infezioni respiratorie) (Cancialosi *et al.*, 2000).

Tale compromissione è comune in numerose malattie neurologiche, in particolare nell'ictus (20-90%), nel morbo di Parkinson (circa 50%), nelle Sclerosi Multiple (33-43%) e nella Sclerosi Laterale Amiotrofica (Gonzalez-Fernández e Daniels, 2008).

La disfagia si verifica nell'85% dei pazienti affetti da Sclerosi Laterale Amiotrofica, in particolare nella forma ad esordio bulbare, nella quale insorge sin dalle fasi più precoci della malattia; la sua presenza aumenta la mortalità dei pazienti di almeno otto volte, contribuendo alla morte del 25.9% di essi.

I sintomi tipici della disfagia nella SLA sono l'aumento della durata del tempo del pasto associato alla fatica nell'alimentarsi, la perdita di peso, episodi di soffocamento e ricorrenti infezioni respiratorie (Carpenter *et al.*, 1978). Nella SLA, la disfagia è causata dalla debolezza o dalla spasticità dei muscoli coinvolti nella deglutizione e innervati da diversi nervi cranici:

**Nervo Trigemino (V)** è un nervo a composizione mista, costituito da fibre



sensitive (quantitativamente prevalenti) e da fibre efferenti motorie. Una sua compromissione può comportare deficit sensitivi e motori dei muscoli masticatori.

**Nervo Faciale (VII)** è costituito da fibre motrici. Una sua lesione può comportare la paralisi faciale periferica.

**Nervo Glossofaringeo (IX)** si tratta di un nervo misto. È coinvolto nella recezione della sensibilità gustativa e riveste un ruolo funzionalmente importante nella gestione e coordinazione dei movimenti di deglutizione. Una sua lesione rende deficitaria specialmente la deglutizione dei solidi.

**Nervo Vago (X)** è un nervo misto, costituito da fibre motrici, vegetative e sensitive. Una sua compromissione comporta disturbi motori del palato molle, della faringe e della laringe.

**Nervo Ipoglosso (XII)** è un nervo esclusivamente motorio. La sua lesione bilaterale si manifesta con una progressiva e globale compromissione della motilità della lingua a cui seguono ingravescenti turbe della deglutizione e della masticazione, funzioni per le quali è indispensabile una corretta ed efficiente motilità linguale (Mutani *et al.*, 2007).

### 1.3 Obiettivi dello studio

Data l'elevata incidenza della disfagia in pazienti affetti da Sclerosi Laterale Amiotrofica e visto il ruolo che questa gioca nell'aumento della loro mortalità, è emersa la necessità di approfondire i seguenti temi.

- La possibile presenza di una correlazione tra la gravità della disfagia e la gravità della compromissione funzionale del paziente.
- L'eventuale correlazione esistente tra la gravità della disfagia e il punteggio totale registrato dai nervi cranici presi sia complessivamente che singolarmente.

- La possibilità di discriminare il tipo di esordio della malattia sulla base dei nervi cranici maggiormente compromessi, cosa che consentirebbe un più rapido e mirato intervento nella cura della patologia.

Si ritiene inoltre necessario verificare l'efficacia del Protocollo di Valutazione Clinica dei Nervi Cranici nel valutare i segni bulbari dei pazienti affetti da SLA. Si tratta di uno strumento, strutturato tra il 2014 e il 2015 presso l'IRCCS Ospedale San Camillo di Venezia, che ha lo scopo di valutare quantitativamente la funzionalità dei cinque nervi cranici coinvolti nella deglutizione (V, VII, IX, X, XII) e di descrivere le possibili implicazioni dei deficit di tali nervi sulla funzione deglutitoria.

Il test diagnostico si basa sulla somma dei punteggi dei nervi cranici. Per i nervi cranici V, VII, X e XII vengono considerati più items, ad ognuno dei quali è assegnato un punteggio compreso tra 0 e 3: il punteggio 0 indica un "deficit grave", il punteggio 1 indica un "deficit medio", il punteggio 2 un "deficit lieve" ed infine il punteggio 3 indica una "performance normale". Invece nella valutazione del IX nervo cranico, il punteggio assegnato è compreso tra 0 e 2.

## 1.4 I dati

Il dataset contiene i dati relativi a 69 pazienti (31 femmine e 38 maschi), che si sono sottoposti a valutazione logopedica presso l'Ambulatorio delle Malattie del Motoneurone della Clinica Neurologica del Dipartimento di Neuroscienze dell'Azienda Ospedaliera di Padova. Sono stati inclusi dal campione i pazienti a cui è stata diagnosticata la Sclerosi Laterale Amiotrofica. Tutti i partecipanti sono stati sottoposti a valutazione motoria tramite la ALSFR scale, eseguita dal neurologo, e al Protocollo di Valutazione Clinica dei Nervi Cranici, eseguito dal logopedista.

Le valutazioni raccolte sono state effettuate tra Aprile 2008 e Luglio 2017.

Le variabili analizzate sono:

- Sesso: variabile dicotomica che indica il sesso del paziente (Maschio (M) o Femmina (F));
- Peso: variabile quantitativa che indica il peso del paziente (misurato in kg);
- Esordio: variabile dicotomica con modalità 1 se l'esordio è di tipo spinale, 2 se l'esordio è di tipo bulbare.
- ALS.Tot: variabile qualitativa, trattabile come quantitativa, che rappresenta la ALSFR scale (*Amyotrophic Lateral Sclerosis Functional Rating Scale - Revised*). Si tratta di uno strumento utilizzato per valutare il grado di compromissione funzionale dei pazienti affetti da SLA. La scala ordinale va a valutare 10 items legati alla vita quotidiana, assegnando a ciascuno un punteggio da 0 (impossibilità di svolgere il compito) a 4 (abilità nella norma). I punteggi vanno poi sommati e determinano un punteggio tra 0 e 40, quest'ultimo rappresenta la migliore prestazione possibile. Maggiore è il valore complessivo registrato da tale variabile, migliore è la qualità della vita dei pazienti affetti da SLA (Cedarbaum *et al.*, 1999).

Gli items sono:

1. Linguaggio (poi trattato singolarmente: ALS.Linguaggio, sottoscala di ALSFR);
2. Salivazione (poi trattato singolarmente: ALS.Salivazione, sottoscala di ALSFR);
3. Deglutizione;
4. Abilità di scrittura a mano (con la mano dominante prima dell'esordio);
5. Capacità di tagliare il cibo e di maneggiare utensili;
6. Capacità di vestirsi e di provvedere all'igiene personale;
7. Capacità di girarsi autonomamente a letto e di aggiustarsi le coperte;

8. Deambulazione;
  9. Capacità di salire le scale;
  10. Respirazione.
- Deglutizione: variabile qualitativa, trattabile come quantitativa, che rappresenta la *Functional Oral Intake Scale* (FOIS), scala ordinale i cui punteggi sono associati al grado di severità della disfagia. Presenta 7 modalità: assume valore 1 quando il paziente è altamente disfagico e non viene alimentato per bocca, mentre assume valore 7 quando non c'è presenza di disfagia.

*Nota:* La valutazione è stata fatta quando i pazienti iniziavano ad avere i primi segni di disfagia, di conseguenza il dataset non contiene alcun valore 7.

Scala FOIS:

NUTRIZIONE ENTERALE (livelli 1-3)

1. Nutrizione Enterale Totale (NET)
2. Nutrizione enterale con minime/sporadiche quantità assunte per via orale
3. Nutrizione enterale supplementare con costanti/consistenti quantità assunte per via orale

NUTRIZIONE ORALE TOTALE (livelli 4-7)

4. Nutrizione orale totale di una sola consistenza
5. Nutrizione orale totale di diverse consistenze che richiedono una speciale preparazione extra-orale
6. Nutrizione orale totale in assenza di speciale preparazione extra-orale, ma con limitazioni a specifici alimenti o liquidi
7. Nutrizione orale totale in assenza di limitazioni.

- Tot.NC: variabile quantitativa ordinata il cui punteggio rappresenta la somma dei valori totali dei vari nervi cranici presi in considerazione. Assume valore 0 quando il paziente è grave, 60 quando non presenta problemi.
- Tot.5, Tot.7, Tot.9, Tot.10 e Tot.12: sono 5 variabili quantitative che rappresentano il valore, rispettivamente, del V, del VII, del IX, del X e del XII nervo cranico.

Tali nervi possono assumere valori rispettivamente 0 (grave) - 12 (normale), 0 (grave) - 12 (normale), 0 (grave) - 6 (normale), 0 (grave) - 12 (normale), 0 (grave) - 15 (normale).

## 1.5 Conclusioni

Nel prossimo capitolo si procederà con le analisi esplorative. L'obiettivo è individuare l'eventuale presenza di relazioni significative tra variabili, in particolare tra la gravità della disfagia e quella della compromissione funzionale del paziente, misurata dalla scala ALSFR, e tra il punteggio totale dei nervi cranici e quello della scala FOIS. Verrà poi condotto lo studio della curva ROC per validare l'accuratezza del test diagnostico basato sulla valutazione dei nervi cranici. Infine, si concluderà con la regressione logistica, condotta al fine di studiare l'effetto che le variabili esplicative hanno su quella dipendente, cioè la gravità della disfagia.



## Capitolo 2

# ANALISI ESPLORATIVE PRELIMINARI

L'analisi esplorativa consente di esaminare i dati in modo che essi rivelino andamenti e relazioni, identifichino caratteristiche particolari di insieme e facilitino la loro descrizione e sintesi (Wayne, 2007).

Le analisi sono svolte con il software statistico RStudio ([www.rstudio.com](http://www.rstudio.com)), considerando un livello di significatività del 5%.

### 2.1 Analisi univariate

Per le variabili quantitative sono state calcolate le classiche misure di sintesi riportate nella *Tabella 2.1*. Per le stesse variabili sono state condotte analisi grafiche (boxplot e istogramma).

La normalità viene valutata tramite il diagramma quantile contro quantile e il test di Shapiro-Wilk.

Analizzando la distribuzione della variabile quantitativa Peso si nota una leggera asimmetria, dovuta forse alla presenza di un valore anomalo all'estremità superiore. Dall'istogramma si vede che la classe modale è (60,70] kg. Il diagramma quantile contro quantile mostra uno scostamento dalla retta di riferimento, particolarmente evidente nella coda superiore (*Figura 2.1*).

L'ipotesi di normalità viene rifiutata ad un livello di significatività del 5% dal test di Shapiro-Wilk ( $SW^{oss} = 0.96$  e  $p - value = 0.04$ ).

Anche la distribuzione di ALS.Tot presenta alcuni valori anomali: 3 pazienti registrano un punteggio particolarmente basso della scala ALS. Dall'istogramma emerge una forte asimmetria: la classe modale è (30, 35]. Tale asimmetria viene poi confermata dal diagramma quantile contro quantile, dove l'allontanamento dalla retta di riferimento è sistematico sia nella coda inferiore che in quella superiore (*Figura 2.2*). Il test di Shapiro-Wilk rifiuta l'ipotesi di normalità ad un livello del 5% ( $SW^{oss} = 0.94$  e  $p - value < 0.01$ ).

Andando a considerare le sotto-scale del linguaggio e della salivazione, notiamo che la variabile ALS.Linguaggio presenta un valore anomalo; la variabile ALS.Salivazione è caratterizzata da una coda sinistra molto lunga (*Figura 2.3*).

Variabile	Minimo	1Qu.	Mediana	Media	3Qu.	Max	sd
Peso	40.00	61.00	70.00	69.14	78.00	114.00	12.20
ALS.Tot	49.00	30.00	32.00	32.68	39.00	46.00	7.39
ALS.Linguaggio	0.00	2.00	3.00	2.54	3.00	4.00	0.95
ALS.Salivazione	0.00	2.00	3.00	2.99	4.00	4.00	1.09
Tot.NC	20.00	41.00	48.00	45.25	53.00	57.00	9.52
Tot.5	4.00	9.00	10.00	10.16	12.00	12.00	1.90
Tot.7	5.00	9.00	10.00	10.00	12.00	12.00	1.76
Tot.9	0.00	4.00	5.00	4.93	6.00	6.00	1.28
Tot.10	1.00	5.00	7.00	6.67	8.00	9.00	1.96
Tot.12	0.00	9.00	12.00	10.84	14.00	15.00	4.12

Tabella 2.1: Statistiche descrittive per le variabili quantitative



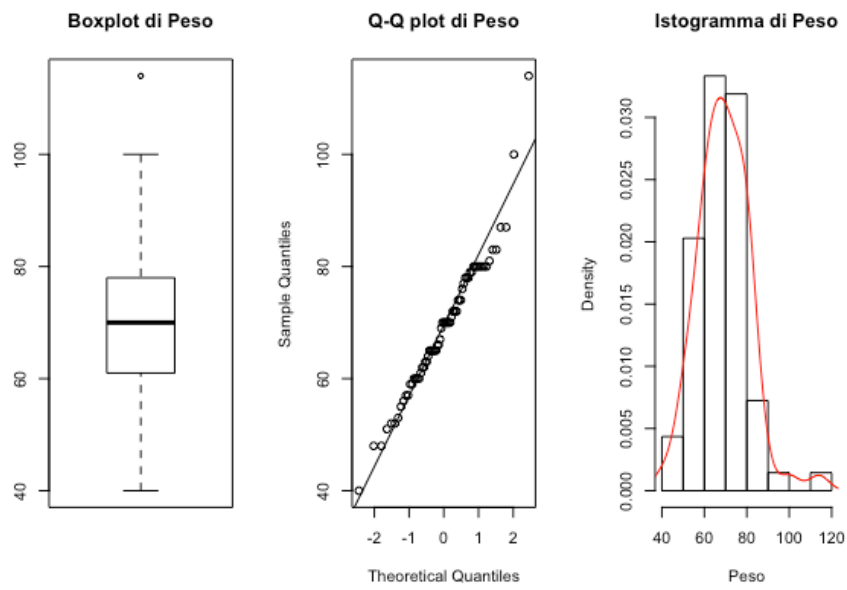


Figura 2.1: Boxplot, Q-Q plot e istogramma per la variabile Peso

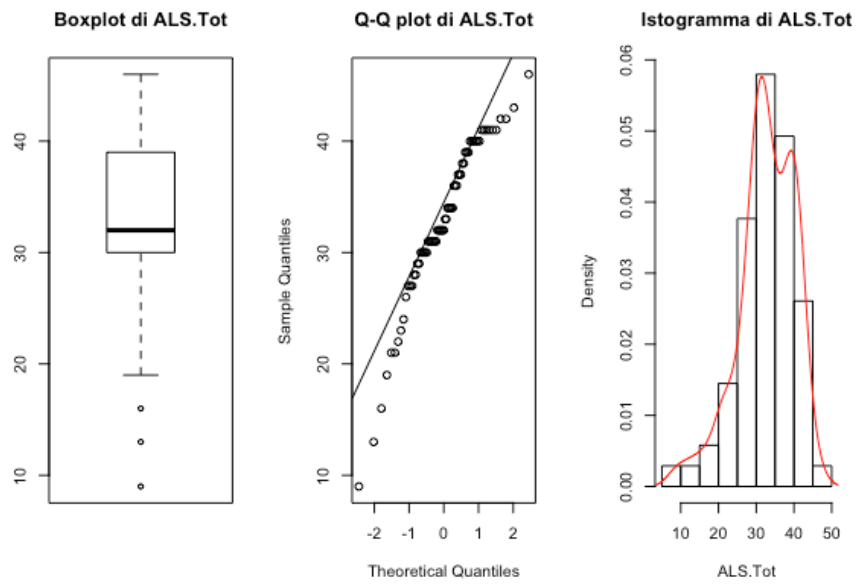


Figura 2.2: Boxplot, Q-Q plot e istogramma per la variabile ALS.Tot

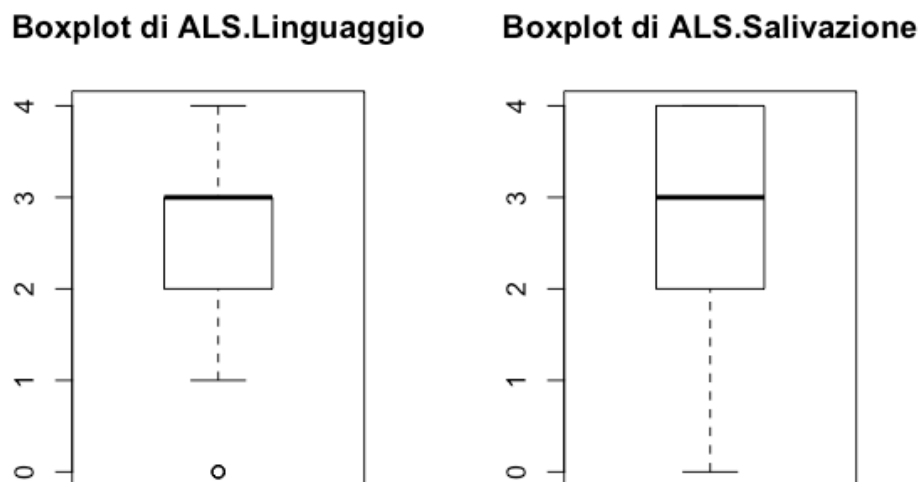


Figura 2.3: Boxplot di ALS.Linguaggio e ALS.Salivazione

L'analisi grafica relativa al punteggio registrato dal totale dei nervi cranici evidenzia una distribuzione asimmetrica, caratterizzata da un valore anomalo e da una coda sinistra molto lunga. Dal diagramma quantile contro quantile emerge un allontanamento dalla retta di riferimento sia nella coda inferiore che in quella superiore (*Figura 2.4*). L'ipotesi di normalità viene rifiutata dal test di Shapiro-Wilk ad un livello di significatività del 5% ( $SW^{oss} = 0.92$  e  $p - value < 0.01$ ).

Dal boxplot in *Figura 2.5* si nota come i nervi maggiormente compromessi siano il Glossofaringeo (IX) e il Vago (X), per i quali i punteggi registrati sono mediamente inferiori. In tutti e 5 i casi le distribuzioni presentano una coda sinistra notevolmente più lunga. Sono presenti alcuni valori anomali per i nervi Trigemino (V), Glossofaringeo (IX) e Ipoglosso (XII).

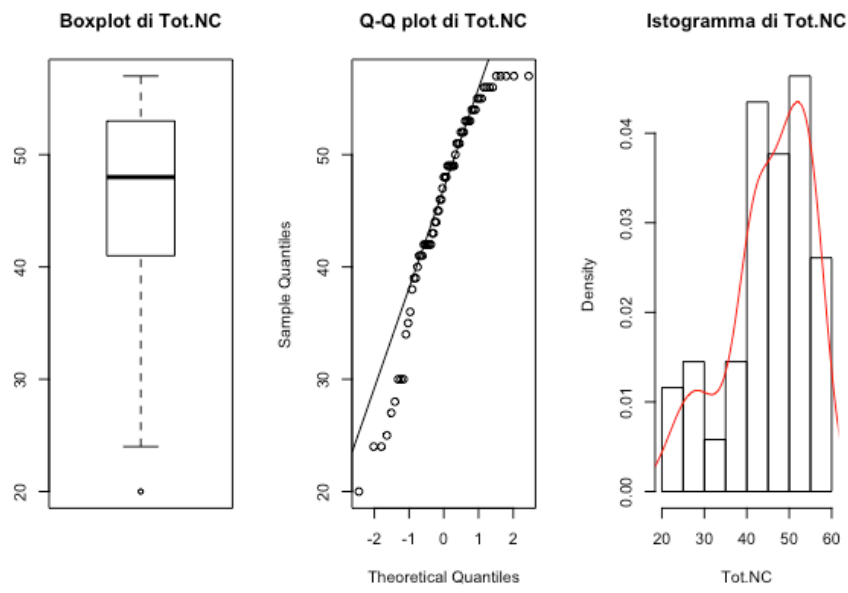


Figura 2.4: Boxplot, Q-Q plot e istogramma per la variabile Tot.NC

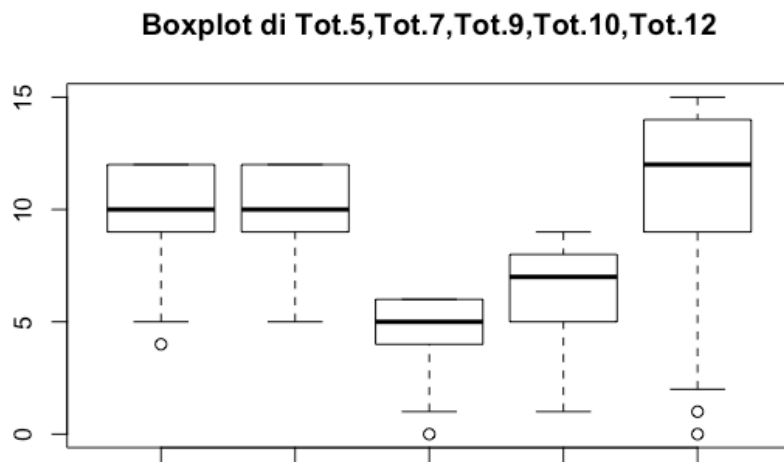


Figura 2.5: Boxplot di Tot.5, Tot.7, Tot.9, Tot.10, Tot.12

Variabile	Modalità	Frequenza assoluta	Frequenza relativa
Sesso	Maschio (1)	38	0.55
	Femmina (0)	31	0.45
Esordio	2: Bulbare (1)	40	0.58
	1: Spinale (0)	29	0.42
Deglutizione	1	1	0.01
	2	8	0.12
	3	7	0.10
	4	28	0.41
	5	18	0.26
	6	7	0.10
	7	0	0.0

Tabella 2.2: Frequenze assolute e relative delle variabili qualitative

Per le variabili qualitative sono state calcolate le frequenze assolute e relative, riportate in *Tabella 2.2*. Il campione è composto da 38 individui di sesso maschile e 31 di sesso femminile (rispettivamente, il 55% e il 45%).

Tra i pazienti, 29 presentano un esordio di tipo spinale (il 42%), per i restanti 40 l'esordio è di tipo bulbare (il 58%). Questo dato è sensato poiché i pazienti considerati sono stati estratti da un campione sottoposto a valutazione logopedica. Bisogna pertanto tenere conto del fatto che la figura del logopedista sia chiamata ad intervenire soprattutto in casi di esordio bulbare; di conseguenza il numero di casi con esordio di tipo 2 sarà maggiore. Inoltre, al momento della raccolta dei dati, 14 pazienti che avevano avuto un esordio spinale manifestavano difficoltà di deglutizione o disfagia, quindi un coinvolgimento bulbare più o meno marcato.

Per quanto riguarda la variabile Deglutizione, il 23% del campione di riferimento ricopre i gradi più gravi della scala FOIS (1-2-3), per i quali è necessaria una nutrizione di tipo enterale; il restante 77% è affetto da lievi problemi disfagici, preservando una nutrizione orale totale. Non è presente

nel campione alcun soggetto sano, privo cioè di problemi disfagici. Questo perché le valutazioni sono state raccolte a partire dai primi segnali di turbe della deglutizione.

## 2.2 Analisi bivariate

Le analisi bivariate indagano le relazioni tra coppie di variabili allo scopo di individuare quelle significative, utilizzando strumenti differenti a seconda della natura delle stesse.

La prima coppia di variabili analizzata riguarda il totale dei nervi cranici e il punteggio della scala ALSFR. Non si nota una relazione lineare tra le due variabili. Sembra che valori più alti della scala ALSFR si addensino in corrispondenza di valori alti del totale di nervi cranici, ma la correlazione di Spearman (0.13) non risulta significativa ( $p - value = 0.30$ ) ad un livello di significatività del 5% (Figura 2.6).

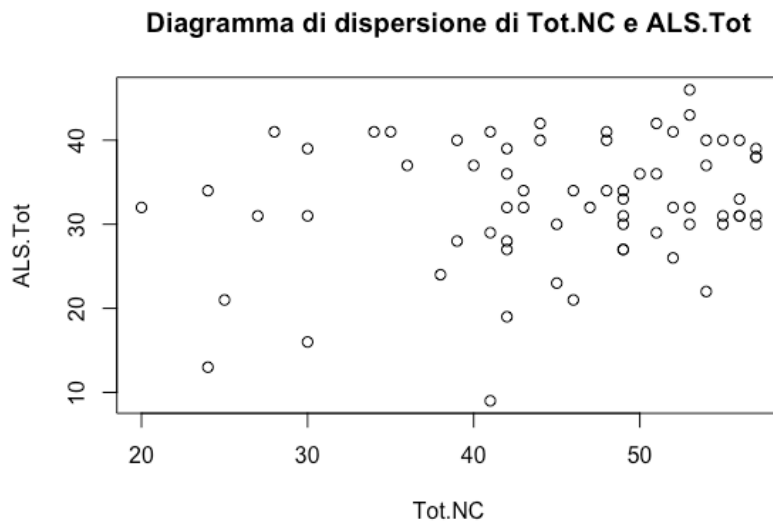


Figura 2.6: Diagramma di dispersione di Tot.NC e ALS.Tot

Anche la relazione tra le variabili Peso - Tot.NC e Peso - ALS.Tot non è di tipo lineare. In entrambi i casi la correlazione di Spearman risulta negativa (rispettivamente -0.09 e -0.13), ma non significativa (rispettivamente  $p - value = 0.46$  e  $p - value = 0.30$ ).

Per l'analisi di coppie di variabili del tipo quantitativa-dicotomica si va a confrontare la distribuzione nei due gruppi individuati dalle modalità della variabile qualitativa dicotomica. Dall'analisi preliminare tramite boxplot si passa a verificare la normalità nei due gruppi tramite l'utilizzo del diagramma quantile contro quantile e del test di Shapiro-Wilk. A seconda che l'ipotesi di normalità sia accettata o meno (nel primo caso si va a valutare anche l'eventuale omoschedasticità), si procede con il test  $t$  a due campioni indipendenti o con il test non-parametrico di Mann-Whitney, rispettivamente.

Dalla *Figura 2.7* appare come il punteggio finale dei nervi cranici sia mediamente inferiore negli individui di sesso maschile. I diagrammi quantile contro quantile mostrano un allontanamento sistematico dalla retta di riferimento in entrambe le code e il test di Shapiro-Wilk rifiuta l'ipotesi di normalità in entrambi i gruppi (per femmina  $SW^{oss} = 0.88$  e  $p - value < 0.01$ ; per maschio  $SW^{oss} = 0.94$  e  $p - value = 0.03$ ). Il test di Mann-Whitney non porta al rifiuto dell'ipotesi di uguaglianza delle mediane nei due gruppi ( $W^{oss} = 635.5$  e  $p - value = 0.58$ ). Il totale dei nervi cranici risulta quindi omogeneo tra pazienti di sesso femminile e maschile.

La compromissione dei nervi cranici appare maggiore nei pazienti con esordio di tipo bulbare (2). La distribuzione di Tot.NC è asimmetrica e la normalità viene rifiutata in entrambi i gruppi dal test di Shapiro-Wilk (per esordio spinale  $SW^{oss} = 0.88$  e  $p - value < 0.01$ ; per esordio bulbare  $SW^{oss} = 0.94$  e  $p - value = 0.05$ ). Il test di Mann-Whitney non porta al rifiuto dell'ipotesi di uguaglianza delle mediane nei due gruppi ( $W^{oss} = 612$  e  $p - value = 0.70$ ) (*Figura 2.8*).

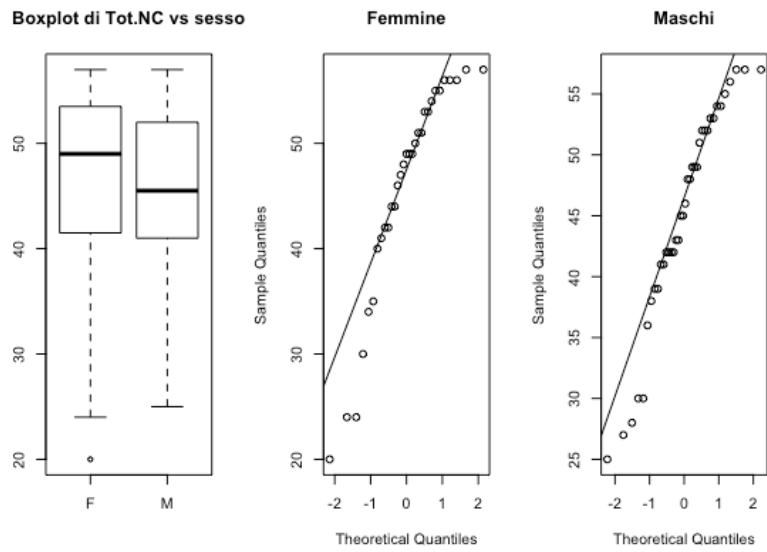


Figura 2.7: Boxplot e diagrammi quantile contro quantile di Tot.NC distintamente per femmine e maschi

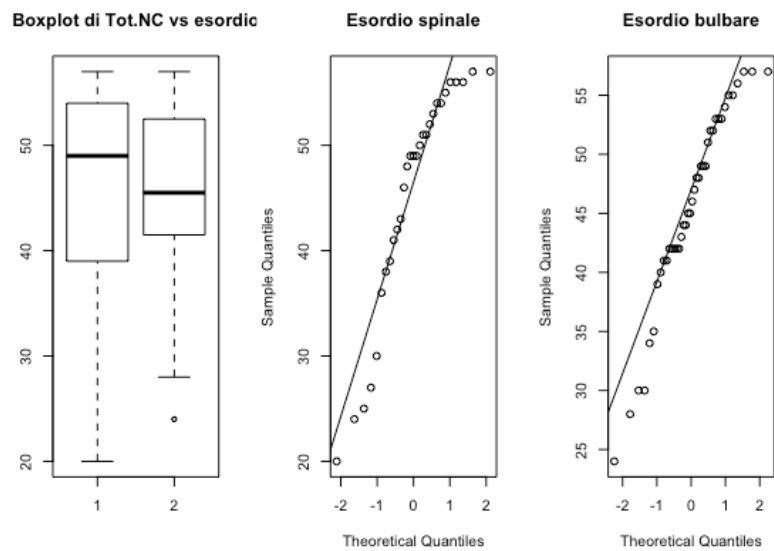


Figura 2.8: Boxplot e diagrammi quantile contro quantile per Tot.NC distintamente nei due gruppi di esordio spinale (1) e bulbare (2)

Dal boxplot in *Figura 2.9* si nota una maggiore compromissione funzionale nei pazienti di sesso maschile. La normalità della distribuzione è accettata solo in uno dei due gruppi (per femmina  $SW^{oss} = 0.90$  e  $p - value < 0.01$ ; per maschio  $SW^{oss} = 0.97$  e  $p - value = 0.29$ ). L'ipotesi di uguaglianza in mediana nei due gruppi non viene rifiutata dal test di Mann-Whitney ad un livello di significatività del 5% ( $W^{oss} = 743.5$  e  $p - value = 0.06$ ).

Dall'analisi della relazione tra le variabili ALS.Tot e esordio emerge un punteggio mediano diverso tra il gruppo con esordio di tipo spinale ed esordio di tipo bulbare. L'ipotesi di normalità viene accettata in entrambi i gruppi (per esordio spinale  $SW^{oss} = 0.95$  e  $p - value = 0.15$ ; per esordio di tipo bulbare  $SW^{oss} = 0.96$  e  $p - value = 0.26$ ), ma si rifiuta l'ipotesi di omoschedasticità ( $p - value < 0.01$ ). Per il confronto tra i due campioni indipendenti si procede dunque con il test di Welch ad alternativa unilaterale ( $t^{oss} = -2.83$  e  $p - value < 0.01$ ), che porta a rifiutare l'ipotesi di uguaglianza in media nei due gruppi. Il punteggio medio della scala ALSFR è significativamente maggiore nei pazienti con esordio di tipo bulbare.

Nell'interpretare tale risultato va tenuto conto del fatto che i pazienti analizzati presentano uno stadio avanzato della malattia, livello al quale ad un esordio di natura spinale si vanno ad aggiungere problemi anche di tipo bulbare.

La *Figura 2.10* mostra come varia il totale dei nervi cranici al variare della gravità della disfagia. È evidente una relazione tra le due variabili: all'aumentare della gravità della disfagia si registra una diminuzione, in mediana, del punteggio dei nervi cranici, che risultano così maggiormente compromessi. La correlazione di Spearman (0.95) risulta infatti significativa ad un livello del 5% ( $p - value < 0.01$ ).



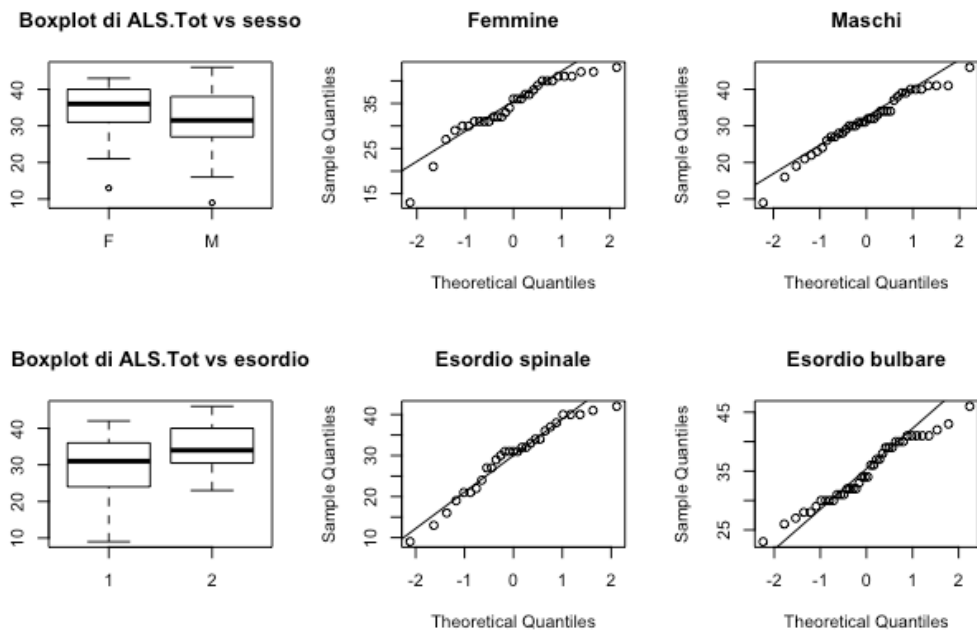


Figura 2.9: Boxplot e Q-Q plots di ALS.Tot distintamente per sesso e tipo di esordio

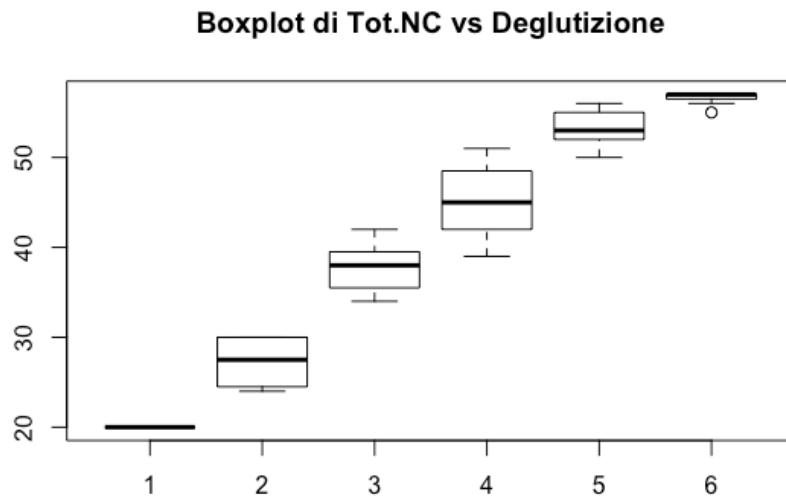


Figura 2.10: Boxplot di Tot.NC per ciascuna modalità assunta dalla variabile Deglutizione

Trattando Deglutizione come variabile numerica, si è stimato un modello di regressione lineare semplice per valutare come varia il punteggio totale dei nervi cranici con quello registrato dalla scala FOIS, che descrive la gravità della disfagia. Dall'analisi si conclude che la dipendenza lineare tra le due variabili è significativa, il modello nullo viene rifiutato ad un livello del 5% ( $p - value < 0.01$ ). L'analisi dei residui non evidenzia nessun problema di eteroschedasticità o non linearità; il test di Shapiro-Wilk non rifiuta l'ipotesi di normalità ( $p - value = 0.11$ ).

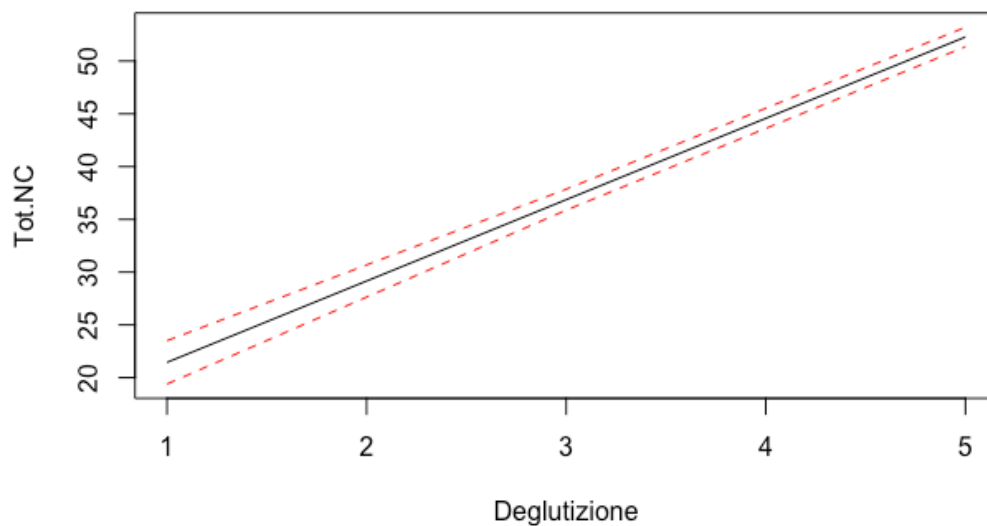


Figura 2.11: Modello di regressione stimato per la relazione tra Deglutizione e Tot.NC (in rosso gli intervalli di previsione per la media della risposta)

La *Tabella 2.3* riporta gli intervalli di previsione relativi al punteggio totale dei nervi cranici in corrispondenza di ciascun valore della scala FOIS. Questo risultato può essere utile per prevedere la gravità della disfagia, conoscendo il range nel quale il valore di Tot.NC rientra.

Deglutizione	Stima	Estremo inferiore	Estremo superiore
1	21.45	19.40	23.51
2	29.16	27.67	30.65
3	36.87	35.88	37.86
4	44.58	43.85	45.30
5	52.28	51.36	53.20
6	59.99	58.60	61.39

Tabella 2.3: Stima, estremo inferiore ed estremo superiore di Tot.NC per ciascun valore assunto da Deglutizione

Per lo studio della dipendenza nei gruppi è stata dicotomizzata la variabile che misura il livello di disfagia. La nuova variabile “deglutizione” assume valore 1 se il valore della scala FOIS (variabile originaria) è inferiore o uguale a 3, cioè se la nutrizione è enterale; viceversa assume valore 0 se la FOIS registra un valore superiore a 3, ovvero nel caso in cui la nutrizione sia di tipo orale.

Dal boxplot in *Figura 2.12* emerge una forte differenza in mediana tra le distribuzioni nei due gruppi. I pazienti del primo gruppo, quelli con lievi problemi disfagici, presentano un punteggio di Tot.NC notevolmente alto, se confrontato con quello registrato dai pazienti del secondo gruppo, quelli cioè con gravi problemi disfagici. Poiché la normalità viene accettata solo per i pazienti a nutrizione enterale (gruppo 1  $SW^{oss} = 0.94$  e  $p - value < 0.01$ ; gruppo 2  $SW^{oss} = 0.96$  e  $p - value = 0.71$ ), si è svolto il test non-parametrico di Mann-Whitney ( $W^{oss}=840$  e  $p-value<0.01$ ). L'ipotesi di uguaglianza in mediana viene rifiutata ad un livello di significatività del 5%. In particolare, il test con alternativa unilaterale conferma quanto già anticipato dal boxplot: il valore mediano di Tot.NC è significativamente inferiore per pazienti con gravi problemi disfagici ( $W^{oss} = 840$  e  $p - value < 0.01$ ).

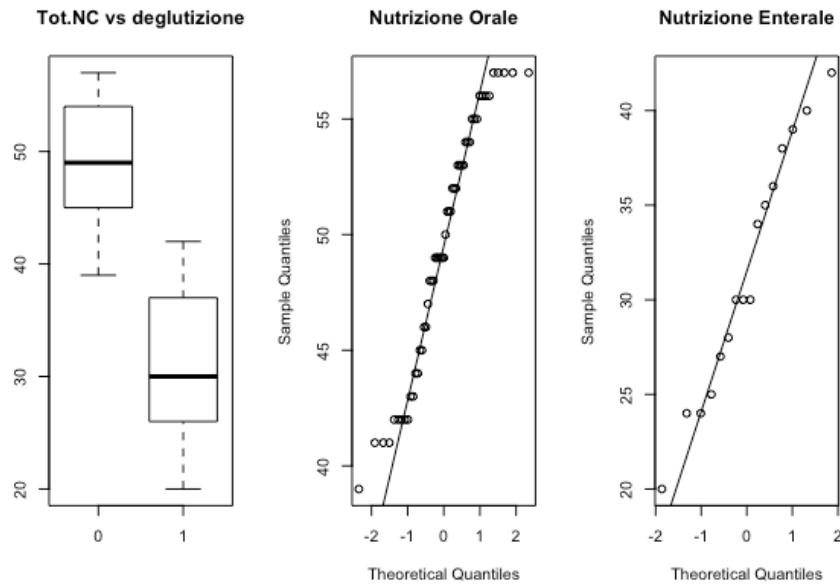


Figura 2.12: Boxplot e diagrammi quantile contro quantile di Tot.NC distintamente nei due gruppi di deglutizione

Per quanto riguarda la variabile peso, non è emersa una differenza significativa tra pazienti a nutrizione orale e pazienti a nutrizione enterale, ad un livello del 5%. L'ipotesi di normalità viene accettata in entrambi i gruppi (gruppo 1  $SW^{oss} = 0.98$  e  $p\text{-value} = 0.43$ ; gruppo 2  $SW^{oss} = 0.93$  e  $p\text{-value} = 0.22$ ), si rifiuta invece l'ipotesi di omoschedasticità ( $p\text{-value} = 0.03$ ). Per il confronto tra i due campioni indipendenti si procede dunque con il test di Welch ( $t^{oss} = -0.58$  e  $p\text{-value} = 0.57$ ), che non porta a rifiutare l'ipotesi di uguaglianza in media nei due gruppi.

Lo stesso tipo di analisi è stata condotta per la variabile ALS.Tot. In questo caso è più difficile individuare una relazione precisa tra il valore della scala ALSFR e la gravità della disfagia. Infatti dal boxplot in *Figura 2.13* non emerge un chiaro andamento di ALS.Tot in funzione delle 6 modalità della variabile Deglutizione. Quanto detto viene confermato dalla correlazione di Spearman (0.18), che non risulta significativa ( $p\text{-value} = 0.13$ ).

**Boxplot di ALS.Tot vs Deglutizione**

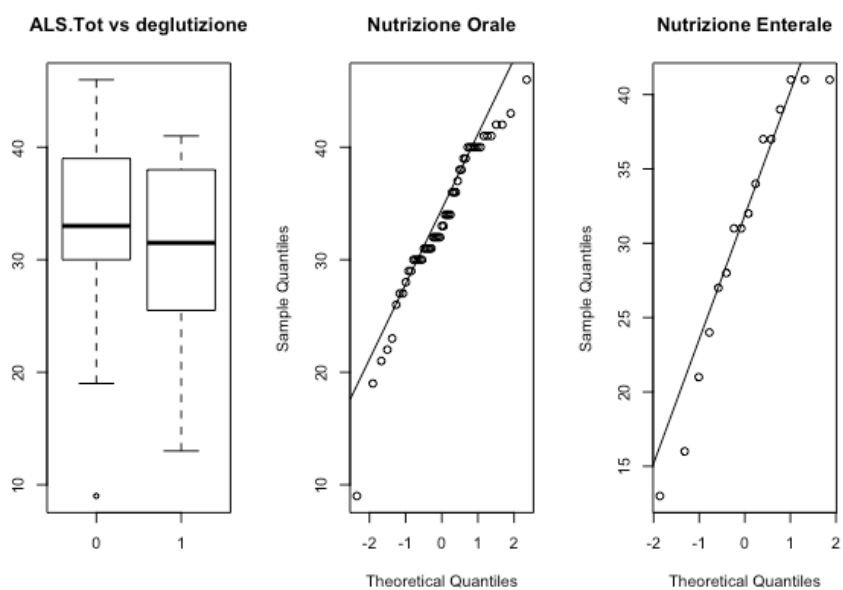
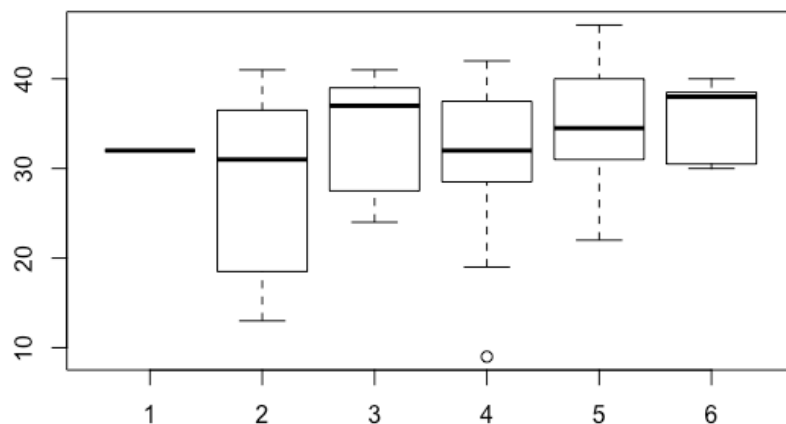


Figura 2.13: Boxplot di ALS.Tot per ciascuna modalità assunta dalla variabile Deglutizione (sopra). Boxplot e Q-Q plots di ALS.Tot distintamente nei due gruppi di deglutizione (sotto)

Anche andando a dicotomizzare il punteggio della scala FOIS non risulta significativa la relazione tra le due variabili. La normalità viene accettata solo in uno dei due gruppi (gruppo 1  $SW^{oss} = 0.94$  e  $p - value = 0.01$ ; gruppo 2  $SW^{oss} = 0.93$  e  $p - value = 0.21$ ). Il test di Mann-Whitney ( $W^{oss} = 368$  e  $p - value = 0.43$ ) non conduce al rifiuto dell'ipotesi di uguaglianza in mediana nei due gruppi.

Tuttavia, se distinguiamo i pazienti in base al tipo di esordio della malattia, la correlazione di Spearman tra la compromissione funzionale e la gravità della disfagia (0.46) diventa significativa per il gruppo ad esordio spinale ( $p - value = 0.01$ ).

Considerando le due sottoscale di ALSFR si ha che il punteggio registrato dall'item relativo al linguaggio si distribuisce in modo significativamente diverso nei due gruppi individuati dalla severità della disfagia; lo stesso non si può dire per l'item relativo alla salivazione. In particolare, per ALS.Linguaggio la normalità viene accettata solo in uno dei due gruppi (gruppo 1  $SW^{oss} = 0.83$  e  $p - value < 0.01$ ; gruppo 2  $SW^{oss} = 0.92$  e  $p - value = 0.18$ ). Il test di Mann-Whitney ( $W^{oss} = 627$  e  $p - value = 0.002$ ) conduce al rifiuto dell'ipotesi di uguaglianza in mediana nei due gruppi ad un livello di significatività del 5%. Invece per quanto riguarda ALS.Salivazione la normalità non viene accettata in nessuno dei due gruppi (gruppo 1  $SW^{oss} = 0.79$  e  $p - value < 0.01$ ; gruppo 2  $SW^{oss} = 0.86$  e  $p - value = 0.01$ ). Il test di Mann-Whitney ( $SW^{oss} = 529$  e  $p - value = 0.12$ ) non porta a rifiutare l'ipotesi di uguaglianza in mediana.

Dall'analisi dei nervi cranici presi singolarmente è emersa una differenza significativa in mediana nei due gruppi individuati dalla gravità della disfagia. In particolare, i nervi cranici maggiormente compromessi si riscontrano nel gruppo 2 di pazienti a nutrizione enterale (gravi problemi disfagici). La normalità viene accettata in uno solo dei due gruppi, per il nervo Glossofaringeo in nessuno dei due ( $p - value < 0.01$ ). Viene pertanto condotto il test non-parametrico di Mann-Whitney per il confronto tra gruppi dei punteggi dei

vari nervi cranici. A livello grafico tale significativa differenza in mediana è confermata dai boxplot rappresentati in *Figura 2.14*.

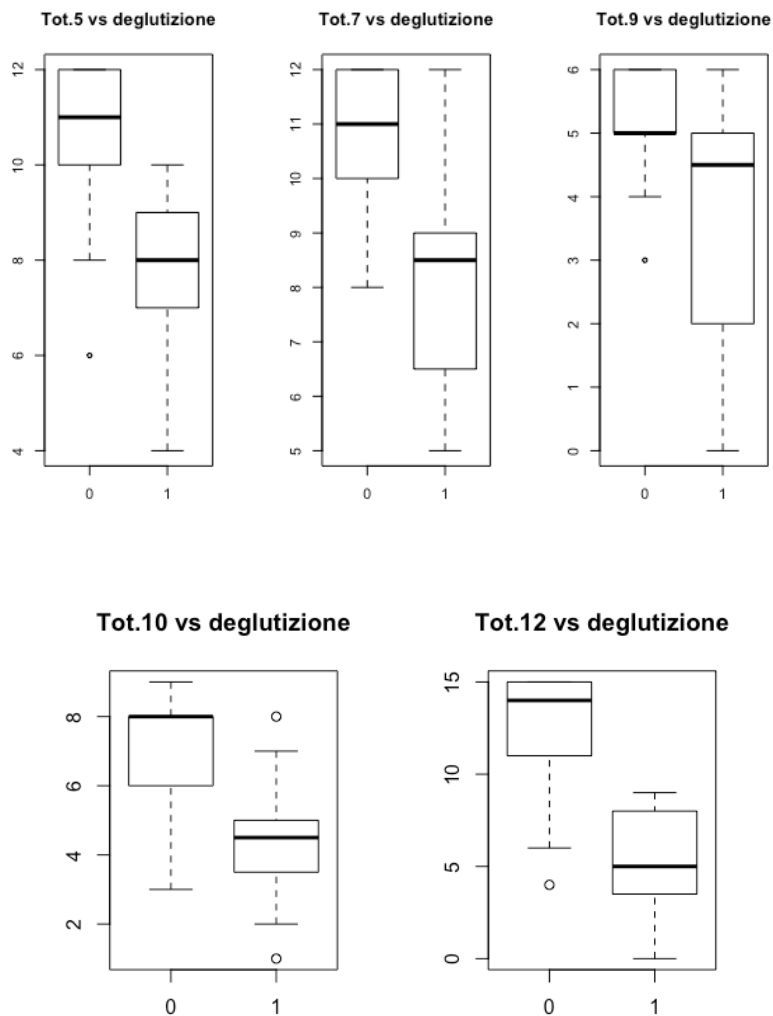


Figura 2.14: Boxplot di Tot.5, Tot.7, Tot.9, Tot.10 e Tot.12 per ciascuna modalità assunta dalla variabile deglutizione

La *Tabella 2.4* riporta il valore osservato e il relativo  $p - value$  di ciascun test condotto per il confronto tra gruppi nel caso dei nervi cranici considerati singolarmente.

	Test Shapiro-Wilk		Test Mann-Whitney	
	SW <sup>oss</sup>	p-value	W <sup>oss</sup>	p-value
Tot.5   deglutizione=0	0.78	<0.01	61	<0.01
Tot.5   deglutizione=1	0.90	0.07		
Tot.7   deglutizione=0	0.87	<0.01	124.5	<0.01
Tot.7   deglutizione=1	0.95	0.50		
Tot.9   deglutizione=0	0.88	<0.01	186	<0.01
Tot.9   deglutizione=1	0.84	<0.01		
Tot.10   deglutizione=0	0.88	<0.01	94	<0.01
Tot.10   deglutizione=1	0.95	0.48		
Tot.12   deglutizione=0	0.85	<0.01	27.5	<0.01
Tot.12   deglutizione=1	0.94	0.35		

Tabella 2.4: Confronto tra gruppi, Test di Shapiro-Wilk e Test di Mann-Whitney



Infine, per l'analisi della coppia di variabili qualitative Deglutizione, considerata nella sua scala originale a 6 modalità, ed esordio, si considera la tabella di contingenza. È stato condotto il test esatto di Fisher, che mostra non esserci dipendenza tra le due variabili ( $p - value = 0.74$ ). Poiché la normalità viene rifiutata in entrambi i gruppi distinti tra esordio di tipo spinale e bulbare (gruppo 1  $SW^{oss} = 0.90$  e  $p - value < 0.01$ ; gruppo 2  $SW^{oss} = 0.90$  e  $p - value < 0.01$ ), si è svolto il test non-parametrico di Mann-Whitney ( $W^{oss} = 598.5$  e  $p - value = 0.82$ ), che non porta a rifiutare l'ipotesi di uguaglianza in mediana.

Analogamente, non è stata riscontrata alcuna dipendenza tra le variabili sesso e Deglutizione ( $p - value = 0.63$ ). La normalità viene rifiutata in entrambi i gruppi (per femmine  $SW^{oss} = 0.91$  e  $p - value = 0.02$ ; per maschi  $SW^{oss} = 0.88$  e  $p - value < 0.01$ ). Il test di Mann-Whitney ( $W^{oss} = 611.5$  e  $p - value = 0.78$ ) non conduce al rifiuto dell'ipotesi di uguaglianza in mediana nei due gruppi.

## 2.3 Conclusioni

Dalle analisi esplorative condotte è emersa una relazione significativa tra il valore totale dei nervi cranici, registrato negli items considerati, e il punteggio della *Functional Oral Intake Scale* (FOIS), associato al grado di severità della disfagia. La dipendenza è di tipo lineare e il modello di regressione stimato consente di disporre di intervalli di previsione, che permettono di individuare lo stadio della disfagia sulla base del punteggio totale dei nervi cranici. Questo risultato conferma l'importanza del Protocollo di Valutazione Clinica dei Nervi Cranici, nella SLA, quale strumento attendibile per la valutazione della funzionalità dei nervi cranici coinvolti nella deglutizione, in grado di fornire un'indicazione affidabile della gravità complessiva della malattia. In particolare, i pazienti che presentano i quadri di disfagia più severa e che, al momento della valutazione o dopo poco tempo, sono passati all'alimentazione enterale (tramite PEG o RIG) hanno ottenuto un punteggio mediamente più basso al Protocollo di Valutazione Clinica, rispetto ai

pazienti ancora in grado di alimentarsi per os. Il risultato del Protocollo di Valutazione Clinica potrebbe perciò essere un iniziale campanello d'allarme per condurre un'analisi più approfondita circa l'apporto nutritivo del paziente e circa la necessità di cambiare modalità di alimentazione.

Per i soli pazienti ad esordio spinale si è inoltre evidenziata una correlazione significativamente positiva tra la gravità della disfagia e quella della compromissione funzionale del paziente, registrata dal punteggio della scala ALSFR. Non risulta, invece, significativa la relazione tra Tot.NC e il tipo di esordio della malattia, cosa che impedisce di poter discriminare l'esordio spinale da quello bulbare sulla base del totale dei nervi cranici.

Anche la variabile sesso non risulta significativa, nel senso che le distribuzioni considerate sono omogenee tra pazienti di sesso maschile e femminile.

# Capitolo 3

## LA CURVA ROC

### 3.1 Introduzione

L'analisi della curva ROC (Curva Operativa Caratteristica o *Receiver Operating Characteristic*) è un metodo statistico che sempre più va diffondendosi in ambito biomedico, in quanto consente di esprimere, rappresentare e valutare l'accuratezza di una procedura diagnostica.

In base alla tipologia di responso fornito, i test si possono dividere in qualitativi e quantitativi. I primi restituiscono un output di tipo dicotomico, i secondi producono risultati sotto forma di variabili numeriche di tipo discrete o continue. Per i test di tipo quantitativo è necessario individuare un valore di soglia che permetta di discriminare i risultati in "positivi" e "negativi". Questo valore prende il nome di *cut-off* (*cut-point*, *threshold*) e può essere identificato tramite l'interpretazione della curva ROC (Bottarelli e Parodi, 2003).

Per ottenere validi risultati attraverso l'uso delle curve ROC è indispensabile che la presenza/assenza di una specifica malattia sia diagnosticata tramite un *golden standard test* (test di riferimento). I test diagnostici ideali (i cosiddetti *golden standard*) discriminano perfettamente i malati dai sani. Tuttavia, quando si considerano i risultati di un particolare test diagnostico su due gruppi, pazienti malati (nel nostro caso gravemente disfagici) e pazienti sani (nel nostro caso lievemente disfagici), è difficile osservare una separazione

netta tra i due gruppi. Infatti, le distribuzioni delle misure del test diagnostico nei due gruppi in genere si sovrappongono parzialmente (Ventura e Racugno, 2017).

## 3.2 Indici sintetici

Il potere diagnostico di un test include la sensibilità, la specificità, il potere predittivo positivo, il potere predittivo negativo e l'accuratezza. Fissata una soglia o *cut-off*  $k$  di discriminazione tra i due gruppi di pazienti, si hanno pazienti malati correttamente classificati come positivi al test (TP=*True Positive* o Veri Positivi), ma anche pazienti malati classificati come negativi (FN=*False Negative* o Falsi Negativi). Inoltre, alcuni pazienti sani appaiono correttamente classificati come negativi al test (TN=*True Negative* o Veri Negativi), ma ci saranno anche pazienti sani classificati come positivi (FP=*False Positive* o Falsi Positivi). Le quattro tipologie TP, FN, TN e FP possono essere rappresentate in una tabella a doppia entrata che contiene il numero di casi classificati nelle quattro modalità possibili. Tale tabella è detta matrice di confusione o tabella di errata classificazione (*Tabella 3.1*).

	Paziente Malato ( $M^+$ )	Paziente Sano ( $M^-$ )	Totale
Test Positivo ( $T^+$ )	TP (Veri Positivi)	FP (Falsi Positivi)	TP+FP
Test Negativo ( $T^-$ )	FN (Falsi Negativi)	TN (Veri Negativi)	FN+TN
Totale	TP+FN	FP+TN	

Tabella 3.1: Matrice di confusione

La validità del test può essere valutata tramite la proporzione di falsi positivi e falsi negativi: tanto più basse saranno tali quote, tanto più il test sarà accurato. L'accuratezza del test è definita come

$$accuratezza = \frac{TP + TN}{TP + FN + FP + TN}.$$

A partire dalla classificazione riportata nella matrice di confusione, si possono poi ottenere due importanti indici sintetici della qualità della performance di un test: la sensibilità e la specificità. La sensibilità è definita come

$$\text{sensibilità} = \frac{TP}{TP + FN}$$

ed esprime la proporzione di Veri Positivi (TP) rispetto al numero totale di pazienti effettivamente malati (TP+FN). Un test diagnostico è sensibile al 100% quando tutti i malati risultano positivi.

La specificità è invece definita come

$$\text{specificità} = \frac{TN}{FP + TN}$$

e misura la proporzione di Veri Negativi (TN) rispetto al numero totale di pazienti sani (FP+TN). Un test diagnostico è specifico al 100% quando tutti i sani risultano negativi. Si noti che se un test ha un'ottima specificità, allora è basso il rischio di pazienti che, pur presentando valori anomali, non sono affetti dalla patologia che si sta ricercando; analogamente, se un test ha un'ottima sensibilità, allora è basso il rischio di pazienti che, pur presentando valori normali, sono comunque affetti dalla patologia.

I valori della sensibilità e della specificità dipendono dalla soglia  $k$  fissata nella classificazione. Quando si fissa un valore elevato di  $k$ , i Veri Positivi e la sensibilità decrescono, mentre la specificità e i Veri Negativi aumentano. Viceversa, se si seleziona un valore basso di  $k$ , i Veri Positivi e la sensibilità aumentano e i Veri Negativi e la specificità diminuiscono.

Nella scelta tra più test si privilegerà quello con maggiore sensibilità o specificità, a seconda che l'obiettivo sia, rispettivamente, individuare il maggiore numero di malati o il maggiore numero di sani.

Altri indici sintetici, calcolabili a partire dalla matrice di confusione, sono il potere predittivo positivo, dato dalla proporzione di pazienti malati ri-

spetto al totale dei positivi al test ( $TP/(TP+FP)$ ), e il potere predittivo negativo, dato dalla proporzione di pazienti sani tra tutti i negativi al test ( $TN/(FN+TN)$ ). Mentre la sensibilità e la specificità sono misure indipendenti dalla prevalenza della malattia, il potere predittivo positivo e quello negativo dipendono strettamente dalla frequenza della malattia di interesse. Si definisce la prevalenza della malattia come il rapporto tra il numero di malati e la totalità della popolazione:

$$prevalenza = \frac{TP + FN}{TP + FN + FP + TN}.$$

Tipicamente questo valore è noto da precedenti studi epidemiologici sulla popolazione.

### 3.3 La curva ROC

Uno strumento per valutare l'accuratezza di un test diagnostico, basato su sensibilità e specificità, è fornito dalla curva ROC. Il modello di classificazione sarebbe ottimale se massimizzasse contemporaneamente sia la sensibilità che la specificità. Questo non è però possibile: elevando il valore della specificità, diminuisce il valore di falsi positivi, ma si aumentano i falsi negativi, il che comporta una diminuzione della sensibilità (esiste un *trade-off* tra i due indici). La relazione tra i precedenti parametri può essere rappresentata riportando, in un sistema di assi cartesiani e per ogni possibile valore di *cut-off*, la proporzione di veri positivi (sensibilità) in ordinata e la proporzione di falsi positivi (1-specificità) in ascissa. L'unione dei punti ottenuti genera una spezzata, la curva ROC.

Un test perfetto dal punto di vista diagnostico (assenza di sovrapposizione tra i due gruppi) è rappresentato da una curva che passa per l'angolo superiore sinistro del sistema di assi cartesiani (100% di specificità, 100% di sensibilità). Diversamente, la bisettrice del quadrato di lati (0,1) corrisponde alla classificazione casuale dei pazienti. Un buon test diagnostico avrà curva

ROC quanto più possibile sopra la bisettrice (Zweig e Campbell, 1993).

Un indicatore di sintesi della curva ROC, che fornisce una misura oggettiva dell'efficacia di un test, cioè della sua capacità di classificazione, è dato dall'area sottesa alla curva, indicata con AUC (*Area Under the Roc Curve*).

Esistono diversi approcci per calcolare quest'area.

Sia  $X$  la variabile che rappresenta la misura nel gruppo dei pazienti sani e  $Y$  quella nel gruppo dei pazienti malati. L'AUC calcolata in base alla probabilità che  $X$  sia dominata da  $Y$  è

$$AUC = P(X < Y).$$

L'AUC può essere stimata sia sotto ipotesi parametriche sulle distribuzioni di  $X$  e  $Y$  sia con metodi non parametrici. L'area sotto la curva può assumere valori compresi tra 0.5 e 1.0 e, al crescere di tale valore, aumenta il potere discriminante del test. Per l'interpretazione del valore stimato dell'AUC si può ricorrere alla classificazione proposta da Swets (1998), regola convenzionale usualmente adottata:

- $AUC = 0.5$ : test non informativo;
- $0.5 < AUC \leq 0.7$ : test poco accurato;
- $0.7 < AUC \leq 0.9$ : test moderatamente accurato;
- $0.9 < AUC < 1.0$ : test altamente accurato;
- $AUC = 1.0$ : test perfetto.

Si possono inoltre calcolare intervalli di confidenza per l'AUC a partire dalla statistica test log-rapporto di verosimiglianza profilo, la cui distribuzione asintotica nulla è  $\chi_1^2$ , o dal test di Wald profilo (Ventura e Racugno, 2017). Infine, per la scelta della soglia necessaria a discriminare tra i due gruppi di pazienti, questa è ottenibile in maniera automatica secondo un opportuno criterio di ottimalità. Per esempio, tramite l'accuratezza, nel senso che si sceglie come soglia il valore che permette di classificare correttamente la

maggior parte dei pazienti. Ancora, indicate con  $Se$  e  $Sp$  la sensibilità e la specificità, due criteri per determinare la soglia ottimale sono il criterio della minima distanza dal punto ottimale  $(0,1)$ :

$$D_{(0,1)} = \sqrt{(1 - Se)^2 + (1 + Sp)^2}$$

e il criterio di Youden o della massima distanza dalla bisettrice:

$$D_{line} = Se + Sp - 1.$$

### 3.4 Applicazione della curva ROC al dataset

Si vuole valutare l'efficacia del test diagnostico sui nervi cranici nel separare individui malati (nel nostro caso gravemente disfagici) da individui sani (lievemente disfagici). Dei 69 pazienti considerati, 53 sono risultati lievemente disfagici (il 77%) e i restanti 16 presentano una disfagia grave (il 23%). Questa rilevazione è stata possibile grazie all'utilizzo di un *golden-test* (test ad alta affidabilità), basato sul punteggio della scala FOIS.

Per valutare la performance del nuovo test diagnostico si ricorre all'analisi ROC. Nel Capitolo 2 è emersa la relazione significativa tra la gravità della disfagia e il punteggio totale dei nervi cranici, infatti la variabile Tot.NC si distribuisce diversamente nei due gruppi a patologia lieve e grave. Tale differenza significativa era stata valutata eseguendo il test non parametrico di Mann-Whitney poiché l'ipotesi di normalità veniva accettata solo in uno dei due gruppi.



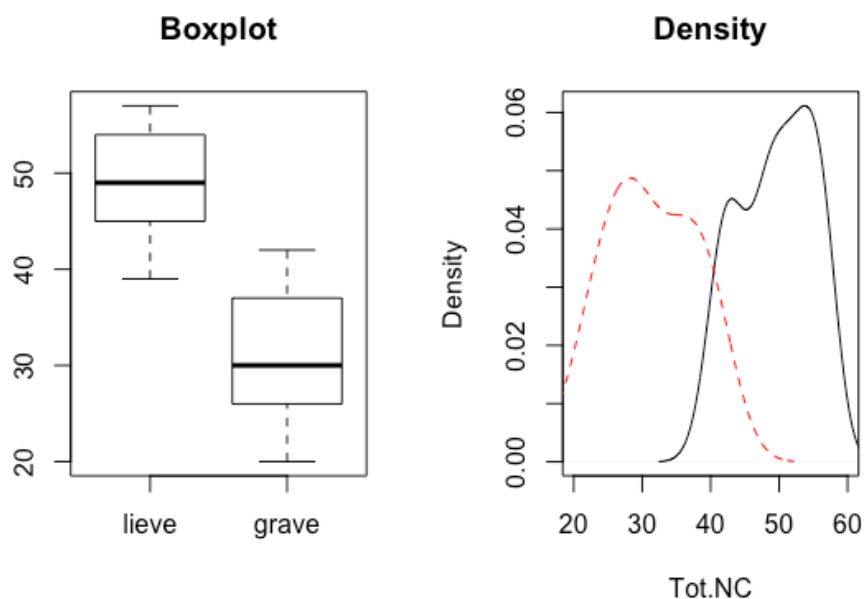


Figura 3.1: Boxplot e densità stimata con il metodo del nucleo di Tot.NC nei pazienti lievemente disfagici (linea nera continua) e nei pazienti gravemente disfagici (linea rossa tratteggiata)

In *Figura 3.1* si può vedere, sia dai due boxplot, sia dalla densità stimata con il metodo del nucleo di Tot.NC nei pazienti sani e malati, una sovrapposizione tra le due distribuzioni. Di conseguenza, per ogni valore di *cut-off* selezionato per discriminare i due gruppi di pazienti, ci saranno pazienti malati classificati come positivi (TP), ma anche pazienti malati classificati come negativi (FN). Analogamente, alcuni pazienti non-malati saranno correttamente classificati come negativi (TN), ma alcuni non-malati saranno classificati come positivi (FP). L'area sotto la curva ROC (*Figura 3.2*) è pari a 0.991, indicando che il test è altamente accurato. L'intervallo di confidenza al 95% per l'AUC è (0.974,1). Sia il criterio di Youden che quello della minima distanza dal punto ottimale (0,1) forniscono come valore soglia un punteggio del totale dei nervi cranici pari a 40.5. Utilizzando il valore soglia trovato, la matrice di confusione risultante è mostrata in *Tabella 3.2*

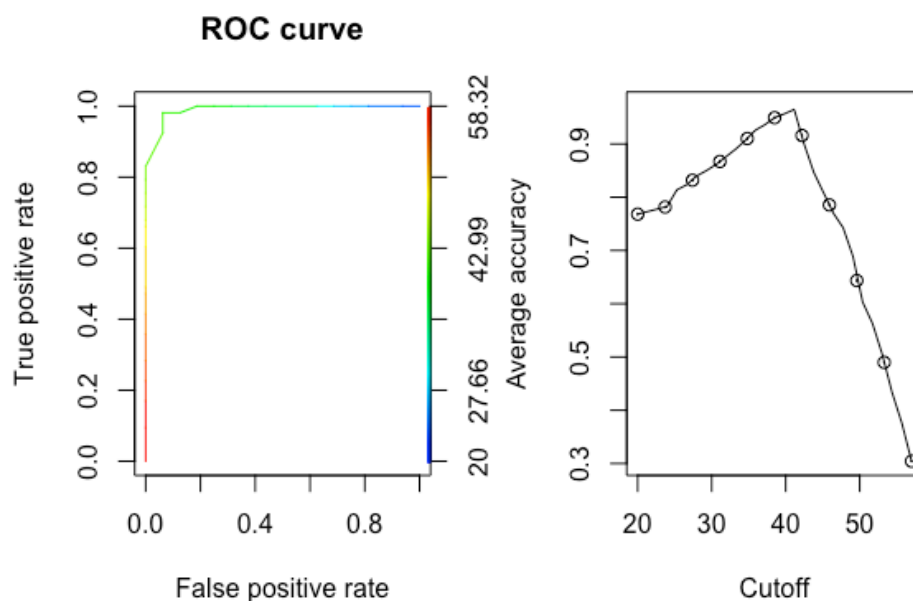


Figura 3.2: Analisi dei livelli di Tot.NC con la curva ROC e l'accuratezza

	Disfagia grave	Disfagia lieve	Totale
Tot.NC $\leq$ 40.5	15	1	16
Tot.NC $>$ 40.5	1	52	53
Totale	16	53	69

Tabella 3.2: Matrice di confusione

La soglia individuata porta ad avere una sensibilità del 94% e una specificità pari al 98%. In particolare, il 6% dei pazienti è falso negativo, viene cioè classificato come sano pur essendo malato; il 2% dei pazienti risulta invece falso positivo, ovvero viene classificato come malato quando è sano. La prevalenza della malattia risulta pari al 23%, l'accuratezza del test diagnostico è del 97%.

Lo stesso tipo di analisi è stata condotta per valutare l'efficacia del test diagnostico basato sui nervi cranici considerati singolarmente.

I risultati sono riportati in *Tabella 3.3*.

	AUC	<i>cut-off</i>	Sensibilità	Specificità	Accuratezza
Tot.5	0.928	9.50	0.88	0.87	0.87
Tot.7	0.853	9.50	0.81	0.75	0.77
Tot.9	0.781	5.50	0.94	0.49	0.60
Tot.10	0.889	5.50	0.81	0.87	0.86
Tot.12	0.968	9.50	1.00	0.87	0.90

Tabella 3.3: Indici sintetici

Risulta che il test diagnostico maggiormente accurato è quello basato su Tot.12. L'area sotto la curva ROC, pari a 0.968 con un intervallo di confidenza (0.933,1), indica che il test è altamente accurato. Questo classifica correttamente il 90% dei pazienti. In particolare, tutti i malati vengono classificati correttamente (*sensibilità* = 100%), mentre il 13% dei pazienti è falso positivo, ovvero risulta malato quando in realtà è sano.

Qualora si volesse privilegiare la specificità, si presentano diverse alternative per la scelta del test migliore: Tot.5, Tot.10 e Tot.12 presentano tutte una specificità del 87%, a diversi livelli di accuratezza. Poco accurato risulta invece il test basato su Tot.9, che classifica correttamente il 60% dei pazienti, con una percentuale di falsi positivi pari al 51%. Dall'intervallo di confidenza per l'AUC (0.661,0.900) si evince che la classificazione dei pazienti tramite il test costruito su tale variabile non è troppo affidabile.

In *Figura 3.3* sono riportate le curve ROC relative a ciascun nervo cranico, con il corrispondente valore di AUC. Viene confermato quanto detto precedentemente: Tot.5 e Tot.12 sono le variabili in grado di discriminare meglio i pazienti.

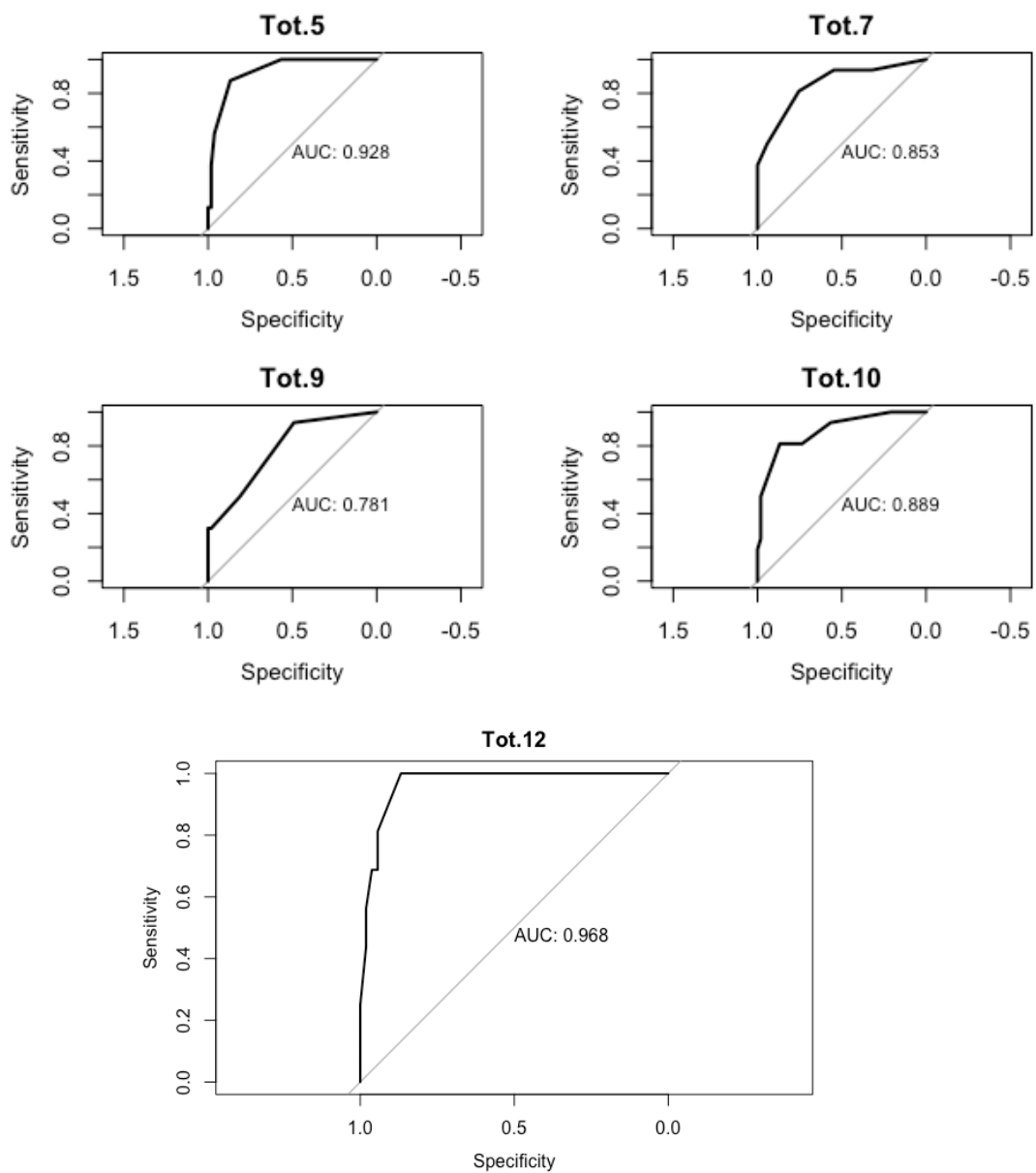


Figura 3.3: Curva ROC e AUC per ciascun nervo cranico

## 3.5 Conclusioni

Lo studio della curva ROC per il test diagnostico sui nervi cranici porta a ritenere tale test altamente accurato. Il 97% dei pazienti sono classificati correttamente. I valori di specificità e di sensibilità sono molto elevati, rispettivamente 98% e 94%, favorendo la corretta classificazione dei sani.

Tutti i criteri impiegati per la determinazione del valore di soglia indicano come *cut-off* un punteggio pari a 40.5 per il totale dei nervi cranici.

Questo vuol dire che il test classifica come gravemente disfagico un paziente con Tot.NC minore, o uguale, a 40.5; viceversa classifica come lievemente disfagico un paziente con Tot.NC maggiore di 40.5.

Per quanto riguarda i test diagnostici sui singoli nervi cranici, si nota una tendenza generale alla corretta classificazione dei malati. Solo con Tot.10 la percentuale di falsi negativi è superiore a quella di falsi positivi, rispettivamente 18% e 13%. Per tutti i singoli nervi la soglia individuata con il criterio di Youden coincide con quella ottenuta minimizzando la distanza dal punto ottimale (0,1). Essa è pari a 9.50 per Tot.5, Tot.7 e Tot.12; 5.50 per Tot.9 e Tot.10. I test più accurati sono quelli basati su Tot.12 e, a seguire, Tot.5, che risultano così le variabili che meglio discriminano i pazienti.



# Capitolo 4

## MODELLI DI REGRESSIONE LOGISTICA

### 4.1 La regressione logistica

La regressione logistica appartiene alla ampia classe dei modelli lineari generalizzati applicabili nelle analisi dove la variabile risposta  $Y$  è non gaussiana ed esprimono la relazione tra la media  $\mu$  della  $Y$  e il predittore lineare  $\eta = \beta_1 x_1 + \dots + \beta_p x_p$  attraverso la funzione di legame  $g(\cdot)$ , rappresentata da

$$g(\mu) = \eta = \beta_1 x_1 + \dots + \beta_p x_p.$$

Il modello di regressione logistica è un modello usualmente impiegato quando la variabile dipendente  $Y$  (variabile risposta) è dicotomica e si vuole modellizzare la media di tale variabile, cioè la probabilità di successo, in funzione di  $p$  variabili esplicative  $x_1, \dots, x_p$ .

La variabile risposta  $Y$  può dunque assumere i valori  $\{0, 1\}$  e segue la distribuzione di Bernoulli, con

$$P(Y = y) = \begin{cases} \pi, & \text{per } y = 1 \\ 1 - \pi, & \text{per } y = 0. \end{cases}$$

La media di  $Y$  è  $E(Y) = \pi = P(Y = 1)$ . Un modello di regressione logistica è quindi una descrizione della dipendenza della probabilità di successo  $\pi$  dai valori che le variabili esplicative  $x_1, \dots, x_p$  assumono. Tale dipendenza può essere formalizzata come

$$g(\pi) = \beta^T x_i = \beta_1 x_1 + \dots + \beta_p x_p,$$

dove  $g(\cdot)$  è la funzione di legame e  $\beta_1, \dots, \beta_p$  sono i parametri di regressione da stimare. Per la funzione di legame  $g(\cdot)$  sono possibili diverse specificazioni. Si ha una regressione logistica quando il legame è una funzione logit (o logistica)

$$\text{logit}(\pi) = \log \frac{\pi}{1 - \pi} = \beta_1 x_1 + \dots + \beta_p x_p = \beta^T x,$$

la cui inversa è

$$\pi = \frac{e^{\beta^T x}}{1 + e^{\beta^T x}}.$$

La funzione di legame logit presenta due vantaggi: è interpretabile in termini di log-rapporto delle probabilità di successo e insuccesso e, negli studi medici, consente di trattare dati raccolti retrospettivamente come se lo studio fosse prospettico. In sintesi, un modello di regressione logistica è specificato dalle seguenti ipotesi:

- $Y_1, \dots, Y_n$  sono variabili indipendenti, con  $Y_i \sim \text{Ber}(\pi_i)$ ;
- il predittore lineare è  $\eta_i = \beta^T x_i$ , con  $x_i = (x_{i1}, \dots, x_{ip})$  riga  $i$ -ma della matrice  $X$  di dimensioni  $(n \times p)$ , non stocastica e con rango pieno;
- la funzione di legame tra  $\eta_i$  e  $\pi_i$  è espressa da

$$g(\pi_i) = \log \frac{\pi}{1 - \pi} = \eta_i = \beta^T x_i.$$

Si può notare come il modello ammetta una specifica forma di eteroschedasticità, in quanto  $\text{Var}(Y_i) = \pi_i(1 - \pi_i)$ ,  $i = 1, \dots, n$ , a differenza del modello lineare che richiede l'assunzione dell'omoschedasticità.



## 4.2 Il problema di separazione e la correzione di Firth

Talvolta, i modelli di regressione logistica incorrono in un problema chiamato fenomeno di separazione, che si presenta principalmente con campioni di piccole dimensioni ed è caratterizzato dalla monotonia della funzione di verosimiglianza. Nonostante tale funzione di verosimiglianza converga, la stima di almeno un parametro non esiste finita, il che comporta l'inesistenza della stima di massima verosimiglianza (Heinze e Schemper, 2002).

Vi sono due tipologie di separazione: quella completa e quella quasi-completa. La prima si ha quando una particolare combinazione dei predittori comporta una previsione perfetta delle osservazioni; la seconda avviene quando quasi tutte le osservazioni vengono classificate correttamente.

Per risolvere il problema di separazione si può condurre una regressione logistica con la correzione di Firth (1993).

Una delle proprietà più importanti dello stimatore di massima verosimiglianza è la non distorsione asintotica. Tuttavia nei campioni finiti lo stimatore è in genere distorto, con distorsione di ordine  $O(n^{-1})$ , che può essere espressa come

$$b(\theta) = \frac{b_1(\theta)}{n} + \frac{b_2(\theta)}{n^2} + \dots,$$

dove  $b_1(\theta), b_2(\theta), \dots$  è un'appropriata sequenza di funzioni, ognuna di ordine  $O(1)$  (Kosmidis, 2014).

Firth (1993) mostra che, nei modelli parametrici regolari, il termine dominante della distorsione asintotica dello stimatore di massima verosimiglianza può essere rimosso tramite un'appropriata modificazione della funzione di punteggio (*score function*), indipendentemente dalla finitezza di  $\hat{\theta}$ .

Si consideri (Firth, 1993) per semplicità di esposizione un modello della famiglia esponenziale con funzione di log-verosimiglianza  $l(\theta) = t\theta - K(\theta)$ , in cui  $p = 1$ ,  $t$  è una statistica sufficiente per il modello e  $K(\theta)$  è una funzione

del parametro. Derivando la funzione di log-verosimiglianza si ottiene

$$U(\theta) = l'(\theta) = t - K'(\theta).$$

In questo caso, la statistica sufficiente  $t$  non influenza la forma di  $U(\theta)$ , ma solo la sua posizione. La distorsione di  $\hat{\theta}$  deriva dalla combinazione di due fattori: la non distorsione della funzione di punteggio ( $E_{\theta}\{U(\theta)\} = 0$  per il vero valore di  $\theta$ ) e la curvatura della funzione di punteggio ( $l''(\theta) \neq 0$ ). Tale distorsione può essere ridotta attraverso l'introduzione di una piccola distorsione nella funzione di punteggio.

In particolare, se  $\hat{\theta}$  è soggetto ad una distorsione positiva (cioè se  $E_{\theta}(\hat{\theta}) > \theta$ ), la funzione di punteggio deve essere spostata verso il basso in ogni punto  $\theta$  di una quantità pari a  $i(\theta)b(\theta)$ , dove  $i(\theta)$  è la matrice di informazione attesa (o informazione di Fisher) e  $b(\theta)$  è la distorsione (*Figura 4.1*).

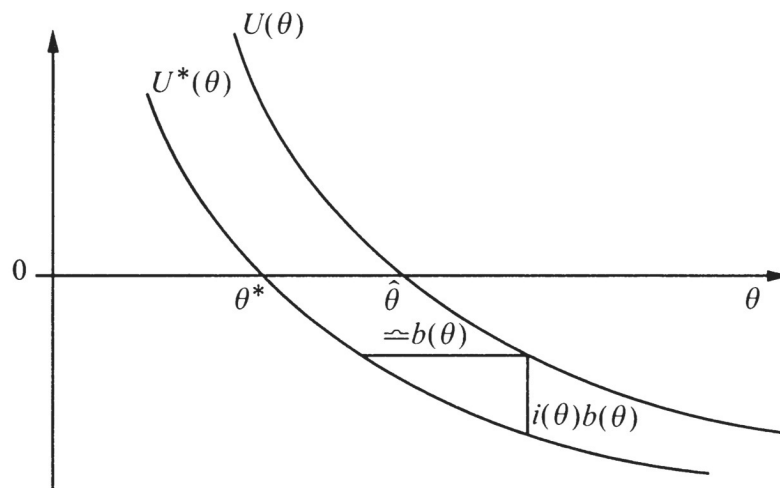


Figura 4.1: Modifica della funzione di punteggio non distorta

La funzione di punteggio così modificata risulta

$$U^*(\theta) = l_*(\theta) - i(\theta)b(\theta).$$

Per calcolare la stima modificata  $\hat{\theta}^*$  si risolve

$$U^*(\theta) = 0$$

e si può dimostrare che  $\hat{\theta}^*$  ha distorsione asintotica di ordine  $O(n^{-2})$ , inferiore a quella di  $\hat{\theta}$ . Le stime secondo il metodo di Firth esistono e sono finite, eliminando completamente il problema di separazione.

### 4.3 Applicazione del metodo di Firth al dataset

Attraverso le regressioni logistiche si vuole studiare l'effetto che le variabili esplicative hanno su quella dipendente. Nel nostro caso la risposta è la variabile dicotomica "deglutizione", che registra la gravità della disfagia per ciascun paziente. In particolare, vale 0 se il paziente è lievemente disfagico (sano) e la nutrizione è di tipo orale; vale 1 se il paziente è gravemente disfagico (malato) e la nutrizione è di tipo enterale.

Nell'applicare il modello logistico al dataset si sono ottenuti come stime dei coefficienti, e dei relativi standard errors, valori particolarmente elevati.

La matrice di confusione ha evidenziato, infatti, un problema di completa separazione. Si è quindi scelto di stimare il modello usando la correzione di Firth. Per evitare un problema di multicollinearità, sono stati stimati due modelli: il primo considerando il punteggio totale dei nervi cranici tra le variabili esplicative; il secondo considerando i punteggi singoli. In entrambi i casi il modello è stato scelto tramite una selezione all'indietro: partendo dal modello con tutte le variabili esplicative si è eliminata, di volta in volta, quella con il minore valore della statistica test del rapporto di verosimiglianza penalizzata, corretta secondo il metodo di Firth.

Il primo modello stimato vuole modellare la probabilità che la disfagia nel paziente sia grave, considerando tra le variabili del modello iniziale il punteggio totale dei nervi cranici. Il modello finale a cui si è giunti ha come unica variabile indipendente Tot.NC. La sua stima è la seguente:

$$\log\left(\frac{\hat{\pi}_i}{1 - \hat{\pi}_i}\right) = 29.401 - 0.743x_{1i},$$

con  $x_{1i} = Tot.NC_i, i = 1, \dots, 69$ .

La statistica test del log-rapporto di verosimiglianza con correzione di Firth porta al rifiuto dell'ipotesi di uguaglianza tra il modello corrente e il modello nullo ( $W^{oss} = 55.38$ ,  $df = 1$ ,  $p - value < 0.01$ ).

	Coef	Std.Error	Chisq-value	p-value
Intercept	29.401139	12.2276026	43.02205	<0.0001
Tot.NC	-0.743121	0.3002319	55.37955	<0.0001

Tabella 4.1: Tabella dei coefficienti del primo modello

Dalla matrice di confusione (*Tabella 4.2*) si ottiene che il modello predice correttamente il 96% delle osservazioni, con una sensibilità pari a  $14/16 = 0.875$  ed una specificità pari a  $52/53 = 0.981$ .

	Test Negativo	Test Positivo
Paziente Sano	52	1
Paziente Malato	2	14

Tabella 4.2: Matrice di confusione per il primo modello

La curva ROC (*Figura 4.2*) è al di sopra della bisettrice, la stima non parametrica dell'AUC è 0.991 con un intervallo di confidenza al 95% pari a (0.974,1), indicando un'elevata adeguatezza del test.

Il parametro  $\beta_1$  è interpretabile come il logaritmo del rapporto di quote conseguente all'aumento unitario della variabile  $x_1$ . Se Tot.NC aumenta di 1, il rapporto di quote è 0.48.

Infine, per completare l'analisi, è stata calcolata la statistica di Hosmer-Lemeshow. L'adattamento tra frequenze attese e stimate è soddisfacente ( $(\chi_{HL}^2)^{oss} = 0.62$  con  $p - value = 0.99$ ).

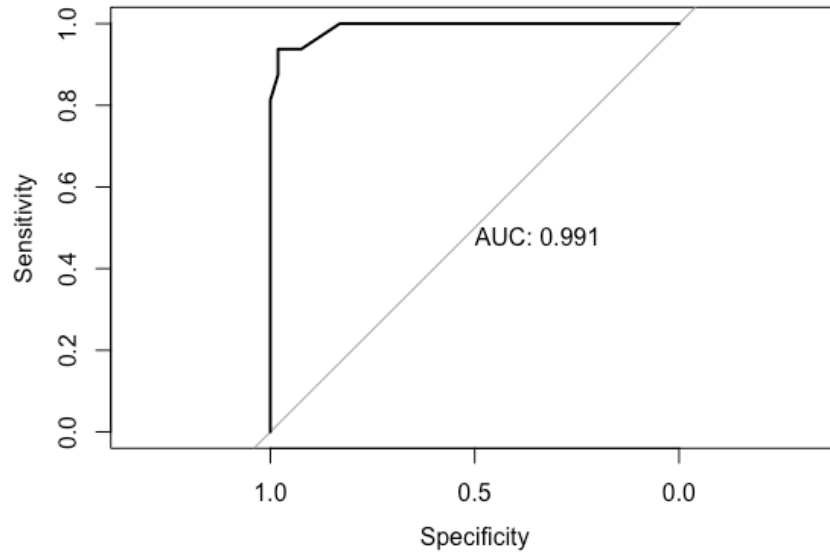


Figura 4.2: Curva ROC per l'accuratezza del primo modello

Il secondo modello stimato considera non più la variabile Tot.NC, ma i singoli punteggi registrati dai nervi cranici (Tot.5, Tot.7, Tot.9, Tot.10, Tot.12). Il modello finale stimato è:

$$\log\left(\frac{\hat{\pi}_i}{1 - \hat{\pi}_i}\right) = 19.838 - 1.495x_{1i} - 0.732x_{2i},$$

con  $x_{1i} = Tot.5_i, i = 1, \dots, 69$  e  $x_{2i} = Tot.12_i, i = 1, \dots, 69$ .

	Coef	Std.Error	Chisq-value	p-value
Intercept	19.8378445	7.3854928	20.74342	<0.0001
Tot.5	-1.4954586	0.6031162	10.03046	<0.0001
Tot.12	-0.7316017	0.2561101	19.85480	<0.0001

Tabella 4.3: Tabella dei coefficienti del secondo modello

Il modello stimato include tra le variabili esplicative quella relativa al punteggio registrato dal V nervo cranico (il nervo Trigemino) e quella riferita al punteggio del XII nervo cranico (il nervo Ipoglosso).

La statistica test del log-rapporto di verosimiglianza con correzione di Firth porta al rifiuto dell'ipotesi di uguaglianza tra il modello corrente e il modello nullo ( $W^{oss} = 51.33$ ,  $df = 2$ ,  $p - value < 0.01$ ).

Dalla matrice di confusione (*Tabella 4.4*) si ottiene che il modello predice correttamente il 93% delle osservazioni, con una sensibilità pari a  $13/16 = 0.813$  ed una specificità pari a  $51/53 = 0.962$ .

	Test Negativo	Test Positivo
Paziente Sano	51	2
Paziente Malato	3	13

Tabella 4.4: Matrice di confusione per il secondo modello

La curva ROC (*Figura 4.3*) è al di sopra della bisettrice, la stima non parametrica dell'AUC è 0.986 con un intervallo di confidenza al 95% pari a (0.966,1), indicando un test altamente accurato.

Il parametro  $\beta_j$  è interpretabile come il logaritmo del rapporto di quote conseguente all'aumento unitario della variabile  $x_j$ . Se Tot.5 aumenta di 1, il rapporto di quote è 0.22; se Tot.12 aumenta di 1, il rapporto di quote è 0.48. Infine, per completare l'analisi, è stata calcolata la statistica di Hosmer-Lemeshow. L'adattamento tra frequenze attese e stimate è soddisfacente ( $(\chi_{HL}^2)^{oss} = 0.61$  con  $p - value = 0.99$ ).

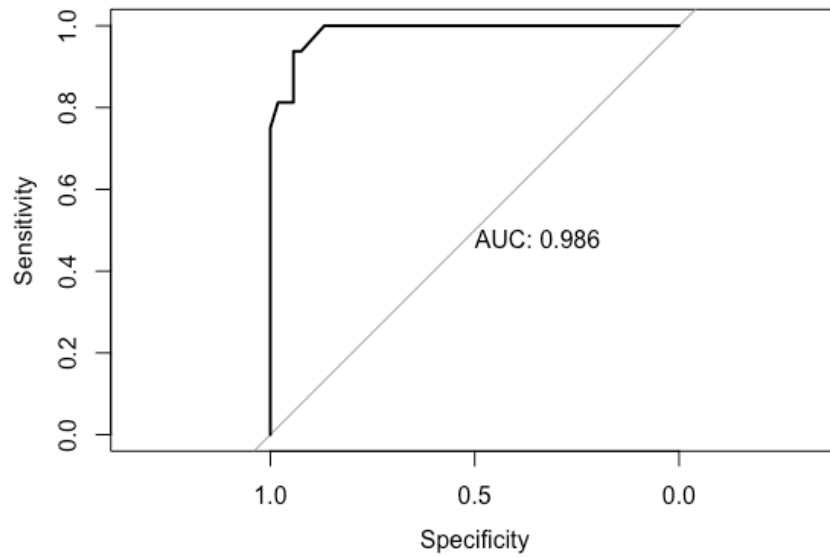


Figura 4.3: Curva ROC per l'accuratezza del secondo modello

## 4.4 Conclusioni

Nel predire la presenza di disfagia, misurata dalla scala FOIS, l'unico contributo è quello dato dalla variabile Tot.NC, considerando il punteggio totale, e Tot.5 insieme a Tot.12, considerando i punteggi singoli. In entrambi i modelli, all'aumentare di tali variabili diminuisce la probabilità di soffrire di disfagia. Si preferisce il primo modello al secondo in quanto presenta una accuratezza lievemente superiore.





# Capitolo 5

## CONCLUSIONI

In questo elaborato si è voluta analizzare la relazione esistente tra la gravità della disfagia e la compromissione funzionale del paziente affetto da SLA, attraverso la valutazione del test diagnostico sui nervi cranici coinvolti nella deglutizione. Nel Capitolo 1, dopo una breve introduzione alla Sclerosi Laterale Amiotrofica e alla disfagia ad essa legata, sono stati presentati i dati. Le unità statistiche corrispondono a 69 pazienti affetti da SLA, con manifestazioni di disfagia a diversi stadi di gravità; le variabili rilevate sono 13 e di varia natura. Esse comprendono alcune caratteristiche demografiche dei pazienti, informazioni legate alla malattia e la performance dei parametri quantitativi rilevati tramite il test diagnostico sui nervi cranici, tramite scala FOIS (per valutare la gravità della disfagia) e, infine, tramite scala ALSFR (per valutare la compromissione funzionale del paziente).

Nel Capitolo 2 sono state condotte le analisi univariate e quelle bivariate, per determinare eventuali relazioni tra variabili. Di particolare rilievo è la relazione significativa emersa tra il punteggio totale dei nervi cranici e quello della scala FOIS, associato al grado di severità della disfagia. La dipendenza osservata è di tipo lineare, è stato quindi stimato un modello di regressione lineare, che consente di disporre di intervalli previsivi al fine di prevedere lo stadio della disfagia sulla base del punteggio individuato. Questo risultato potrebbe consentire un più rapido e mirato intervento nella cura della patologia. Successivamente è stata dicotomizzata la variabile relativa al punteggio

della scala FOIS, individuando due gruppi di pazienti: quelli gravemente disfagici a nutrizione enterale e quelli lievemente disfagici a nutrizione orale. Dall'analisi dei gruppi è emerso che sia il totale dei nervi cranici sia i nervi considerati singolarmente sono in grado di discriminare i pazienti. Al contrario, il tipo di esordio della malattia, il peso e il sesso del paziente non sembrano influenzare le distribuzioni considerate, tuttavia emerge una correlazione significativamente positiva tra la gravità della disfagia e quella della compromissione funzionale del paziente per i soli pazienti ad esordio spinale. Nel Capitolo 3 è stata considerata la curva ROC considerando come variabile criterio il totale dei nervi cranici e ottenendo come soglia un punteggio pari a 40.5. Dallo studio condotto vi è evidenza nel ritenere che il test diagnostico sui nervi cranici sia altamente accurato, classificando correttamente il 97% dei pazienti con una sensibilità ed una specificità pari, rispettivamente, al 94% e 98%. Per quanto riguarda i test diagnostici costruiti sui nervi cranici presi singolarmente, emerge che le variabili che meglio discriminano i pazienti sono Tot.5 e Tot.12. Entrambi i test hanno una specificità del 87%; la sensibilità è massima per il test basato su Tot.12 (100%), invece per quello basato su Tot.5 è pari a 88%.

Infine, nel Capitolo 4 sono stati stimati due modelli di regressione logistica con lo scopo di determinare le variabili in grado di influenzare la presenza di disfagia. In particolare, considerando il totale dei nervi cranici, si ha che l'unica variabile significativa è Tot.NC. Invece, se si considerano i nervi cranici singolarmente, risulta che la disfagia dipende da Tot.5 e da Tot.12, confermando quanto ottenuto dall'analisi della curva ROC. Difatti una compromissione del nervo Trigemino (V) può comportare deficit sensitivi e motori dei muscoli masticatori; una lesione del nervo Ipoglosso (XII) si manifesta con una progressiva e globale compromissione della motilità della lingua, indispensabile per una corretta masticazione e deglutizione.

Concludendo, i risultati ottenuti confermano l'importanza del Protocollo di Valutazione Clinica dei Nervi Cranici quale strumento attendibile nell'indicare la gravità complessiva della disfagia nella SLA, sulla base della valutazione dei nervi cranici coinvolti nella deglutizione.

# Bibliografia

- [1] Bottarelli, E. e Parodi, S. (2003). *Un approccio per la valutazione della validità dei test diagnostici: le curve R.O.C. (Receiver Operating Characteristic)*. Ann. Fac. Medic. Vet. di Parma, 23, 49-68.
- [2] Cancialosi, P., Gonella, M.L., Unnia, L. e Buratti, M.G. (2000). *Manuale operativo di fisiopatologia della deglutizione*. Omega edizioni, Torino.
- [3] Carpenter, R.J., McDonald, T.J. e Howard, F.M. (1978). *The otolaryngologic presentation of amyotrophic lateral sclerosis*. Otolaryngology, 86, 479-484.
- [4] Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B. e Nakanishi, A. (1999). *The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III)*. Journal of Neurological Science, 169, 13-21.
- [5] Dodds, W.J., Stewart, E.T. e Logemann, J.A. (1990). *Physiology and radiology of the normal oral and pharyngeal phases of swallowing*. AJR Am J Roentgenol, 154, 953-963.
- [6] Firth, D. (1993). *Bias reduction of maximum likelihood estimates*. Biometrika, 80, 27-38.
- [7] Gonzalez-Fernández, M. e Daniels, S.K. (2008). *Dysphagia in stroke and neurologic disease*. Physical Medicine and Rehabilitation Clinics of North America, 19, 867-888.

- [8] Heinze, G. e Schemper, M. (2002). *A solution to the problem of separation in logistic regression*. *Statistics in Medicine*, 21, 2409-2419.
- [9] Kosmidis, I. (2014). *Bias in parametric estimation: reduction and useful side-effects*. *WIRE Computational Statistic*, 6, 185-196.
- [10] Mutani, R., Lopiano, L., Durelli, L., Mauro, A. e Chiò, A. (2011). *Il Bergamini di neurologia*. Cortina.
- [11] Rozzanigo, G. (2017). Tesi di Laurea. *Valutazione dei nervi cranici e della disfagia in pazienti affetti da Sclerosi Laterale Amiotrofica*. Dipartimento di Neuroscienze DNS, Università di Padova.
- [12] Ventura, L. e Racugno, W. (2017). *Biostatistica. Casi di studio in R*. Egea.
- [13] Wayne, D. (2007). *Biostatistica*. Edises.
- [14] Wijesekera, L.C. e Leigh, P.N. (2009). *Amiyotrophic Lateral Sclerosis*. *Orphanet Journal of Rare Disease*, 4, 3-25.
- [15] Zweig, H.H. e Campbell, G. (1993). *Receiver Operating Characteristic (ROC) plots: a fundamental evolution tool in medicine*. *Clin. Chem.*, 39, 561-77.
- [16] Dispense didattiche di "Statistica Medica". A.A. 2017-2018. A cura della Prof.ssa Laura Ventura.
- [17] Materiale didattico di "Modelli Statistici 2". A.A. 2017-2018. A cura dei Proff. Alessandra Salvan, Nicola Sartori e Luigi Pace.