Matricola:

⓪	⓪	⓪	⓪	⓪	⓪
①	①	①	①	①	①
②	②	②	②	②	②
③	③	③	③	③	③
④	④	④	④	④	④
⑤	⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨	⑨

*Istruzioni: riempire **completamente** le bolle con le cifre del numero di matricola (una cifra per colonna); nella parte sotto del foglio, riempire **completamente** le bolle con le risposte alle domande a scelta multipla. Per riempire, usare penna o matita nera, colorando tutto l'interno e cercando di non uscire dal bordo. Non sono ammesse correzioni, dato che il foglio verrà analizzato da un computer.*

Cognome:............................**Nome**:............................**Firma**:............................

Segnare le risposte delle domande a scelta multipla

(1) Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

(2) Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

(3) Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

(4) Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

(5) Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ

Domande a scelta multipla

(1) Che cosa rappresenta il livello di significatività α di un test?

- ☐ *(a)* [=] La massima probabilità di rifiutare H_0 quando è vera.
- ☐ *(b)* La minima probabilità di rifiutare H_0 quando è vera.
- ☐ *(c)* La massima probabilità di accettare H_0 quando è falsa.
- ☐ *(d)* La minima probabilità di accettare H_0 quando è falsa.
- ☐ *(e)* La massima probabilità di rifiutare H_0 quando è falsa.

(2) In un test d'ipotesi per la media a varianza nota con $H_0 : \mu \leq \mu_0$ e $H_1 : \mu > \mu_0$ si rifiuta H_0 ad un livello $\alpha = 0.05$ in corrispondenza ad un campione di ampiezza n. Si prenda ora un campione di ampiezza $m > n$ tale che $\overline{x}_m = \overline{x}_n$; cosa succede al $P - value\ \overline{\alpha}$? (Sugg: scrivere la regione di rifiuto e ricordarsi che $\phi(x) < 1/2$ se e solo se $x < 0$.)

- ☐ *(a)* non cambia.
- ☐ *(b)* [=] diminuisce. *(*) Supponiamo $\alpha < 1/2$ (non necessariamente pari a 0.05). La regione di rifiuto è $(\bar{x}_n - \mu_0)\sqrt{n}/\sigma > q_{1-\alpha}$ ed il P-value è $\bar{\alpha} = 1 - \phi((\bar{x}_n - \mu_0)\sqrt{n}/\sigma)$. Essendo $\alpha < 1/2$ si ha $q_\alpha > 0$ e, poiché si rifiuta, $\bar{x}_n > \mu_0$. Quindi se $n < m$, allora $(\bar{x}_m - \mu_0)\sqrt{m}/\sigma > (\bar{x}_m - \mu_0)\sqrt{n}/\sigma = (\bar{x}_n - \mu_0)\sqrt{n}/\sigma$. Il P-value di conseguenza diminuisce.*
- ☐ *(c)* non si può determinare con esattezza, perché non è noto il segno di $\bar{x}_m - \mu_0$.
- ☐ *(d)* dipende dal segno di $\overline{x}_n$.
- ☐ *(e)* aumenta.

(3) Si consideri un test di adattamento per un campione di ampiezza n. Cosa è sempre vero in caso n aumenti mentre il livello di confidenza e le frequenze relative osservate rimangono le stesse?

- ☐ *(a)* Il P-value aumenta.
- ☐ *(b)* La regione di rifiuto diventa più piccola.
- ☐ *(c)* Nessuna delle altre risposte è vera.
- ☐ *(d)* [=] Da un certo valore di n in poi H_0 verrà rifiutata.
- ☐ *(e)* Da un certo valore di n in poi H_0 verrà accettata.

(4) Si supponga di avere un campione $\{(x_i, y_i)\}_{i=1}^n$ di ampiezza n dove $y_i > 0$ per ogni i. Si sospetta una relazione del tipo $y = a \cdot b^x$ con $a, b > 0$. Come procediamo per verificare l'attendibilità di tale ipotesi?

- ☐ *(a)* Regressione lineare di y contro e^x.
- ☐ *(b)* [=] Regressione lineare di $\ln(y)$ contro x. *(*) Chiaramente ciò su può fare se e solo se $y_i > 0$ per ogni i oppure $y_i < 0$ per ogni i. In tal caso il segno di a coincide con quello comune delle $\{y_i\}_i$. A questo punto, determinato il segno, si ha $\ln(|y_i|) = x\ln(b) + \ln(|a|)$. In questo caso abbiamo supposto $y_i > 0$ pertanto se esiste una*

soluzione deve soddisfare $a > 0$ e quindi i moduli non sono necessari. Naturalmente la base del logaritmo non è importante, qui si può utilizzare $\log_c$ per qualsiasi $c \in (0, +\infty) \setminus \{1\}$.

- ☐ *(c)* Regressione lineare di e^y contro x.
- ☐ *(d)* Regressione lineare di y contro x.
- ☐ *(e)* Regressione lineare di y contro $ln(|x|)$.

(5) Sia X_1, X_2, X_3, X_4 un campione casuale di ampiezza 4 con $X_i \sim B(8, p)$, $i = 1, 2, 3, 4$ dove p è un parametro incognito. Data la variabile aleatoria T_1 definita da:

$$T_1 = X_1 + X_2 + X_3 + X_4 - 32p,$$

si ha che

- ☐ *(a)* Nessuna delle altre risposte è sempre corretta.
- ☐ *(b)* $E[T_1] = 8p$ per ogni p.
- ☐ *(c)* [=] $T_2 = X_4\,(T_1 + 32p)$ è uno stimatore (non necessariamente corretto) di p.
- ☐ *(d)* T_1 è uno stimatore (non necessariamente corretto) di p.
- ☐ *(e)* $T_2 = T_1 + 16p$ è uno stimatore (non necessariamente corretto) di p.