Formulario per l'esame di Statistica Allievi INF, TEL. AA 08/09 Docente: Ilenia Epiifani
Probabilità

Variabili aleatorie discrete

Nome simbolo parametri	Densità discreta $p_X(k) = \mathrm{P}[X = k]$	F.g.m. $m_X(t) = \mathrm{E}\left(\mathrm{e}^{tX}\right)$	Media	Varianza
bernoulliana $X \sim \mathbf{Be}(p) = \mathbf{Bi}(1, p)$ $0 < p < 1$	$p^k(1-p)^{1-k}$ $k = 0, 1$	$(1-p) + p\mathrm{e}^t$	p	$p(1-p)$
binomiale $X \sim \mathbf{Bi}(n, p)$ $0 < p < 1$, $n = 1, 2, \ldots$	$\binom{n}{k} p^k(1-p)^{n-k}$ $k = 0, 1, \ldots, n$	$\left(1 - p + p\mathrm{e}^t\right)^n$	np	$np(1-p)$
ipergeometrica $(b + r, r, n)$ $b, r \geq 0, n \geq 1$ interi $n \leq b + r$	$\frac{\binom{r}{k}\binom{b}{n-k}}{\binom{b+r}{n}}$ $\max\{0, n-b\} \leq k \leq \min\{r, n\}$ k intero		$\frac{nr}{b+r}$	$\frac{nrb}{(b+r)^2}\left(1 - \frac{n-1}{b+r-1}\right)$
geometrica (p) $0 < p < 1$	$p_X(k) = p(1-p)^{k-1}$ $k = 1, 2, \ldots$	$\dfrac{p\mathrm{e}^t}{1-(1-p)\mathrm{e}^t}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
di Poisson $X \sim \mathcal{P}(\lambda)$ $\lambda > 0$	$\dfrac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$ $k = 0, 1, \ldots$	$\exp\left[\lambda\left(\mathrm{e}^t - 1\right)\right]$	λ	λ
uniforme su $1, 2, \ldots, n$ $n \geq 1$ intero	$\dfrac{1}{n}$ $k = 1, 2, \ldots, n$	$\dfrac{\mathrm{e}^t\left(1 - \mathrm{e}^{nt}\right)}{n\left(1 - \mathrm{e}^t\right)}$	$\dfrac{n+1}{2}$	$\dfrac{n^2-1}{12}$

Variabili aleatorie continue

Nome simbolo parametri	Densità continua $f_X(x) = F'_X(x)$ $\forall x$ t.c. F'_X esiste	F.g.m.	Media	Varianza
uniforme $X \sim \mathcal{U}(a, b)$ $a < b$	$\dfrac{1}{b-a}\mathbf{1}_{(a,b)}(x)$	$\begin{cases} \dfrac{\mathrm{e}^{bt} - \mathrm{e}^{at}}{(b-a)\,t} & \text{se } t \neq 0 \\ 1 & \text{se } t = 0 \end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
normale [gaussiana] $X \sim \mathrm{N}(\mu, \sigma^2)$ $\mu \in \mathbb{R}$, $\sigma > 0$	$\dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$	$\exp\left(\mu t + \dfrac{1}{2}\sigma^2 t^2\right)$	μ	σ^2
lognormale $\mu \in \mathbb{R}$, $\sigma > 0$	$\dfrac{1}{x\sqrt{2\pi\sigma^2}}\exp\left[-\dfrac{(\ln x - \mu)^2}{2\sigma^2}\right]$ $(x > 0)$		$\mathrm{e}^{\mu + \sigma^2/2}$	$\mathrm{e}^{2\mu + \sigma^2}\left(\mathrm{e}^{\sigma^2} - 1\right)$
gamma $X \sim \Gamma(a, \beta)$ $a, \beta > 0$	$\dfrac{1/\beta^a}{\Gamma(a)}x^{a-1}\mathrm{e}^{-x/\beta}\mathbf{1}_{(0,+\infty)}(x)$	$\dfrac{1}{(1-\beta t)^a}$ $\forall t < 1/\beta$	$a\beta$	$a\beta^2$
esponenziale $X \sim \mathcal{E}(\beta)$ $\beta > 0$	$\mathcal{E}(\beta) = \Gamma(1, \beta)$	$\dfrac{1}{1-\beta t}$ $\forall t < 1/\beta$	β	β^2
chiquadro $X \sim \chi^2_n$ $n \geq 1$, intero	$\chi^2_n = \Gamma(n/2\,, 2)$	$\left(\dfrac{1}{1-2t}\right)^{n/2}$ $\forall\, t < \frac{1}{2}$	n	$2n$
Weibull $\alpha, \beta > 0$	$\dfrac{\alpha}{\beta}x^{\alpha-1}\,\mathrm{e}^{-x^\alpha/\beta}\mathbf{1}_{(0,\infty)}(x)$		$\Gamma(1 + 1/\alpha)\beta^{1/\alpha}$	$\beta^{2/\alpha}[\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)]$
t di Student $X \sim t_n$ $n = 1, \ldots$	$\dfrac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\dfrac{n}{2}\right)}\left(1 + \dfrac{x^2}{n}\right)^{-(n+1)/2}$		0 se $n > 1$	$\dfrac{n}{n-2}$ se $n > 2$

Integrale gamma: $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}\,\mathrm{e}^{-x}\,\mathrm{d}x$ e $\Gamma(a+1) = a\Gamma(a)$ $\forall a > 0$

Vettore gaussiano bivariato $(X, Y)^T \sim \mathcal{N}(\mu_X, \mu_Y, \sigma^2_X, \sigma^2_Y, \varrho)$ $\quad [\varrho = \dfrac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}]$

$$f_{X\ Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\varrho^2}}\mathrm{e}^{-\frac{1}{2(1-\varrho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\varrho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}$$

$\mathrm{Var}(X) = \mathrm{E}(X - \mathrm{E}(X))^2 = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$
$\mathrm{Cov}(X, Y) = \mathrm{E}((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$

f.d.r bidimensionale continua: $F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(s, t)\; ds\; dt$
f.d.r bidimensionale discreta: $F_{X,Y}(x, y) = \sum_{s \leq x}\sum_{t \leq y} p_{X,Y}(s, t)$
Densità marginali di (X, Y) continuo: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\; dy$
Densità marginali di (X, Y) discreto: $p_X(x) = \sum_{y_k} p_{X,Y}(x, y_k)$

1 Test di ipotesi sulla media di una popolazione gaussiana

$(x_1, \ldots, x_n)$ = realizzazione campionaria di $X_1, \ldots, X_n$ *i.i.d.* $\sim N(\mu, \sigma^2)$.

σ^2 nota [z-test]:

$\mathbf{H_0}$	$\mathbf{H_1}$	**Si rifiuta $\mathbf{H_0}$ se**	**p-value**
$\mu=\mu_0$ $\mu=\mu_0$ $\mu\leq\mu_0$	$\mu=\mu_1$ *con* $\mu_0<\mu_1$ $\mu>\mu_0$ $\mu>\mu_0$	$\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \geq z_{1-\alpha}$	$1-\Phi\left(\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right)$
$\mu=\mu_0$ $\mu=\mu_0$ $\mu\geq\mu_0$	$\mu=\mu_1$ *con* $\mu_0>\mu_1$ $\mu<\mu_0$ $\mu<\mu_0$	$\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \leq -z_{1-\alpha}$	$\Phi\left(\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right)$
$\mu=\mu_0$	$\mu\neq\mu_0$	$\dfrac{\lvert\bar{x}-\mu_0\rvert}{\sigma/\sqrt{n}} \geq z_{1-\frac{\alpha}{2}}$	$2\left[1-\Phi\left(\dfrac{\lvert\bar{x}-\mu_0\rvert}{\sigma/\sqrt{n}}\right)\right]$

σ^2 incognita [t-test]:

$\mathbf{H_0}$	$\mathbf{H_1}$	**Si rifiuta $\mathbf{H_0}$ se**	**p-value**
$\mu=\mu_0$ $\mu=\mu_0$ $\mu\leq\mu_0$	$\mu=\mu_1$ *con* $\mu_0<\mu_1$ $\mu>\mu_0$ $\mu>\mu_0$	$\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}} \geq t_{n-1}(1-\alpha)$	$1-P\left(T_{n-1}\leq\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\right)$
$\mu=\mu_0$ $\mu=\mu_0$ $\mu\geq\mu_0$	$\mu=\mu_1$ *con* $\mu_0>\mu_1$ $\mu<\mu_0$ $\mu<\mu_0$	$\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}} \leq -t_{n-1}(1-\alpha)$	$P\left(T_{n-1}\leq\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\right)$
$\mu=\mu_0$	$\mu\neq\mu_0$	$\dfrac{\lvert\bar{x}-\mu_0\rvert}{s/\sqrt{n}} \geq t_{n-1}(1-\dfrac{\alpha}{2})$	$2\left[1-P\left(T_{n-1}\leq\dfrac{\lvert\bar{x}-\mu_0\rvert}{s/\sqrt{n}}\right)\right]$

$\bar{x}$ = media campionaria di $x_1, \ldots, x_n$
s^2 = varianza campionaria di $x_1, \ldots, x_n$
Φ = f.d.r. $N(0,1)$ e z_p t.c. $\Phi(z_p) = p$
$T_{n-1} \sim t$ di student con $n-1$ gradi di libertà e $t_{n-1}(p)$ t.c. $P(T_{n-1} \leq t_{n-1}(p)) = p$.

2 χ^2-test sulla varianza di una popolazione gaussiana

μ **nota:**

$\mathbf{H_0}$	$\mathbf{H_1}$	**Si rifiuta $\mathbf{H_0}$ se**	**p-value**
$\sigma^2=\sigma_0^2$ $\sigma^2=\sigma_0^2$ $\sigma^2\leq\sigma_0^2$	$\sigma^2=\sigma_1^2$ *con* $\sigma_0^2<\sigma_1^2$ $\sigma^2>\sigma_0^2$ $\sigma^2>\sigma_0^2$	$\dfrac{ns_0^2}{\sigma_0^2} \geq \chi_n^2(1-\alpha)$	$1-F_n\left(\dfrac{ns_0^2}{\sigma_0^2}\right)$
$\sigma^2=\sigma_0^2$ $\sigma^2=\sigma_0^2$ $\sigma^2\geq\sigma_0^2$	$\sigma^2=\sigma_1^2$ *con* $\sigma_0^2>\sigma_1^2$ $\sigma^2<\sigma_0^2$ $\sigma^2<\sigma_0^2$	$\dfrac{ns_0^2}{\sigma_0^2} \leq \chi_n^2(\alpha)$	$F_n\left(\dfrac{ns_0^2}{\sigma_0^2}\right)$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\dfrac{ns_0^2}{\sigma_0^2} \geq \chi_n^2(1-\dfrac{\alpha}{2})$ oppure $\dfrac{ns_0^2}{\sigma_0^2} \leq \chi_n^2(\dfrac{\alpha}{2})$	$2\min\{p_1,p_2\}$ dove $p_1=F_n\left(\frac{ns_0^2}{\sigma_0^2}\right)$ e $p_2=1-p_1$

$s_0^2 = (1/n)\sum_{j=1}^n (x_j-\mu)^2$.

μ **incognita:**

$\mathbf{H_0}$	$\mathbf{H_1}$	**Si rifiuta $\mathbf{H_0}$ se**	**p-value**
$\sigma^2=\sigma_0^2$ $\sigma^2=\sigma_0^2$ $\sigma^2\leq\sigma_0^2$	$\sigma^2=\sigma_1^2$ *con* $\sigma_0^2<\sigma_1^2$ $\sigma^2>\sigma_0^2$ $\sigma^2>\sigma_0^2$	$\dfrac{(n-1)s^2}{\sigma_0^2} \geq \chi_{n-1}^2(1-\alpha)$	$1-F_{n-1}\left(\dfrac{(n-1)s^2}{\sigma_0^2}\right)$
$\sigma^2=\sigma_0^2$ $\sigma^2=\sigma_0^2$ $\sigma^2\geq\sigma_0^2$	$\sigma^2=\sigma_1^2$ *con* $\sigma_0^2>\sigma_1^2$ $\sigma^2<\sigma_0^2$ $\sigma^2<\sigma_0^2$	$\dfrac{(n-1)s^2}{\sigma_0^2} \leq \chi_{n-1}^2(\alpha)$	$F_{n-1}\left(\dfrac{(n-1)s^2}{\sigma_0^2}\right)$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\dfrac{(n-1)s^2}{\sigma_0^2} \geq \chi_{n-1}^2(1-\dfrac{\alpha}{2})$ o $\dfrac{(n-1)s^2}{\sigma_0^2} \leq \chi_{n-1}^2(\dfrac{\alpha}{2})$	$2\min\{p_1,p_2\}$ dove $p_1=F_{n-1}\left(\frac{(n-1)s^2}{\sigma_0^2}\right)$ e $p_2=1-p_1$

F_n = funzione di ripartizione chi-quadro con n gradi di libertà e $\chi_n^2(p)$ t.c. $F_n(\chi_n^2(p)) = p$.

3 Test per il confronto di medie di due popolazioni gaussiane

3.1 Caso di campioni indipendenti:

$(x_1, \dots, x_m)$ = realizzazione di $\boldsymbol{X} = X_1, \dots, X_m$ *i.i.d.* $\sim N(\mu_X, \sigma_X^2)$,
$(y_1, \dots, y_n)$ = realizzazione di $\boldsymbol{Y} = Y_1, \dots, Y_n$ *i.i.d.* $\sim N(\mu_Y, \sigma_Y^2)$ e $\boldsymbol{X}, \boldsymbol{Y}$ indipendenti.

σ_X^2, σ_Y^2 note [z-test]:

$\mathbf{H_0}$	$\mathbf{H_1}$	**Si rifiuta $\mathbf{H_0}$ se**	**p-value**
$\mu_X = \mu_Y + \Delta$ $\mu_X \leq \mu_Y + \Delta$	$\mu_X > \mu_Y + \Delta$ $\mu_X > \mu_Y + \Delta$	$\dfrac{\bar{x} - \bar{y} - \Delta}{\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}} \geq z_{1-\alpha}$	$1 - \Phi\left(\dfrac{\bar{x} - \bar{y} - \Delta}{\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}}\right)$
$\mu_X = \mu_Y + \Delta$ $\mu_X \geq \mu_Y + \Delta$	$\mu_X < \mu_Y + \Delta$ $\mu_X < \mu_Y + \Delta$	$\dfrac{\bar{x} - \bar{y} - \Delta}{\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}} \leq -z_{1-\alpha}$	$\Phi\left(\dfrac{\bar{x} - \bar{y} - \Delta}{\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}}\right)$
$\mu_X = \mu_Y + \Delta$	$\mu_X \neq \mu_Y + \Delta$	$\dfrac{\lvert\bar{x} - \bar{y} - \Delta\rvert}{\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}} \geq z_{1-\frac{\alpha}{2}}$	$2\left[1 - \Phi\left(\dfrac{\lvert\bar{x} - \bar{y} - \Delta\rvert}{\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}}\right)\right]$

$\sigma_X^2 = \sigma_Y^2$ incognite ma uguali [t-test]:

$\mathbf{H_0}$	$\mathbf{H_1}$	**Si rifiuta $\mathbf{H_0}$ se**	**p-value**
$\mu_X = \mu_Y + \Delta$ $\mu_X \leq \mu_Y + \Delta$	$\mu_X > \mu_Y + \Delta$ $\mu_X > \mu_Y + \Delta$	$\dfrac{\bar{x} - \bar{y} - \Delta}{s_p\sqrt{1/m + 1/n}} \geq t_{m+n-2}(1-\alpha)$	$1 - P\left(T_{m+n-2} \leq \dfrac{\bar{x} - \bar{y} - \Delta}{s_p\sqrt{1/m + 1/n}}\right)$
$\mu_X = \mu_Y + \Delta$ $\mu_X \geq \mu_Y + \Delta$	$\mu_X < \mu_Y + \Delta$ $\mu_X < \mu_Y + \Delta$	$\dfrac{\bar{x} - \bar{y} - \Delta}{s_p\sqrt{1/m + 1/n}} \leq -t_{m+n-2}(1-\alpha)$	$P\left(T_{m+n-2} \leq \dfrac{\bar{x} - \bar{y} - \Delta}{s_p\sqrt{1/m + 1/n}}\right)$
$\mu_X = \mu_Y + \Delta$	$\mu_X \neq \mu_Y + \Delta$	$\dfrac{\lvert\bar{x} - \bar{y} - \Delta\rvert}{s_p\sqrt{1/m + 1/n}} \geq t_{m+n-2}(1 - \frac{\alpha}{2})$	$2 - 2P\left(T_{m+n-2} \leq \dfrac{\lvert\bar{x} - \bar{y} - \Delta\rvert}{s_p\sqrt{1/m + 1/n}}\right)$

$s_p^2 = \frac{s_X^2(m-1) + s_Y^2(n-1)}{m+n-2}$ con s_X^2 = varianza campionaria di $x_1, \dots, x_m$ e s_Y^2 = varianza campionaria di $y_1, \dots, y_n$,
$T_{m+n-2} \sim t$ di student con $m + n - 2$ gradi di libertà.

3.2 Caso di campioni accoppiati:

$(X_1, Y_1), \dots, (X_n, Y_n)$ *i.i.d.* $\sim N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \varrho\sigma_X\sigma_Y \\ \varrho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}\right)$ e $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \varrho$ incogniti. Per i problemi di verifica di ipotesi:

– $H_0 : \mu_X \leq \mu_Y + \Delta$ contro $H_1 : \mu_X > \mu_Y + \Delta$
– $H_0 : \mu_X \geq \mu_Y + \Delta$ contro $H_1 : \mu_X < \mu_Y + \Delta$
– $H_0 : \mu_X = \mu_Y + \Delta$ contro $H_1 : \mu_X \neq \mu_Y + \Delta$

svolgere opportuno t-test usando il campione $X_1 - Y_1, \dots, X_n - Y_n$.

4 F-test per il confronto di varianze di due popolazioni gaussiane

$(x_1,\dots,x_m)$ = realizzazione di $\boldsymbol{X} = X_1,\dots,X_m$ *i.i.d.* $\sim N(\mu_X,\sigma_X^2)$,
$(y_1,\dots,y_n)$ = realizzazione di $\boldsymbol{Y} = Y_1,\dots,Y_n$ *i.i.d.* $\sim N(\mu_Y,\sigma_Y^2)$ e $\boldsymbol{X},\boldsymbol{Y}$ indipendenti.

μ_X, μ_Y **note:**

$\mathbf{H_0}$	$\mathbf{H_1}$	**Si rifiuta $\mathbf{H_0}$ se**	**p-value**
$\sigma_X^2 \le \sigma_Y^2$	$\sigma_X^2 > \sigma_Y^2$	$\dfrac{s_{0X}^2}{s_{0Y}^2} \ge F_{m,n}(1-\alpha)$	$1 - P\left(F_{m,n} \le \dfrac{s_{0X}^2}{s_{0Y}^2}\right)$
$\sigma_X^2 \ge \sigma_Y^2$	$\sigma_X^2 < \sigma_Y^2$	$\dfrac{s_{0X}^2}{s_{0Y}^2} \le F_{m,n}(\alpha)$	$P\left(F_{m,n} \le \dfrac{s_{0X}^2}{s_{0Y}^2}\right)$
$\sigma_X^2 = \sigma_Y^2$	$\sigma_X^2 \ne \sigma_Y^2$	$s_{0X}^2/s_{0Y}^2 \ge F_{m,n}(1-\alpha/2)$ oppure $s_{0X}^2/s_{0Y}^2 \le F_{m,n}(\alpha/2)$	$2\min\{p_1,p_2\}$ dove $p_1 = P\left(F_{m,n} \le s_{0X}^2/s_{0Y}^2\right)$ e $p_2 = 1-p_1$

$s_{0X}^2 := \frac{\sum_{j=1}^m (x_j-\mu_X)^2}{m}$ e $s_{0Y}^2 := \frac{\sum_{j=1}^n (y_j-\mu_Y)^2}{n}$.
$F_{a,b}$ = v.a. avente densità di Fisher con (a,b) gradi di libertà e $F_{a,b}(p)$ t.c. $P(F_{a,b} \le F_{a,b}(p)) = p$.

μ_X, μ_Y **incognite:**

$\mathbf{H_0}$	$\mathbf{H_1}$	**Si rifiuta $\mathbf{H_0}$ se**	**p-value**
$\sigma_X^2 \le \sigma_Y^2$	$\sigma_X^2 > \sigma_Y^2$	$\dfrac{s_X^2}{s_Y^2} \ge F_{m-1,n-1}(1-\alpha)$	$1 - P\left(F_{m-1,n-1} \le \dfrac{s_X^2}{s_Y^2}\right)$
$\sigma_X^2 \ge \sigma_Y^2$	$\sigma_X^2 < \sigma_Y^2$	$\dfrac{s_X^2}{s_Y^2} \le F_{m-1,n-1}(\alpha)$	$P\left(F_{m-1,n-1} \le \dfrac{s_X^2}{s_Y^2}\right)$
$\sigma_X^2 = \sigma_Y^2$	$\sigma_X^2 \ne \sigma_Y^2$	$s_X^2/s_Y^2 \ge F_{m-1,n-1}(1-\alpha/2)$ oppure $s_X^2/s_Y^2 \le F_{m-1,n-1}(\alpha/2)$	$2\min\{p_1,p_2\}$ dove $p_1 = P\left(F_{m-1,n-1} \le s_X^2/s_Y^2\right)$ e $p_2 = 1-p_1$