Probabilità e Processi Stocastici (455AA)

Lezione 10

Dario Trevisan – *https://web.dm.unipi.it/trevisan*

22/10/2025

Variabili aleatorie gaussiane

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri
- l'analisi delle componenti principali (PCA) e una giustificazione tramite opportune variabili gaussiane

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri
- l'analisi delle componenti principali (PCA) e una giustificazione tramite opportune variabili gaussiane
- qualche rudimento della regressione, in particolare il metodo dei minimi quadrati, giustificato tramite opportune ipotesi di gaussianità

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri
- l'analisi delle componenti principali (PCA) e una giustificazione tramite opportune variabili gaussiane
- qualche rudimento della regressione, in particolare il metodo dei minimi quadrati, giustificato tramite opportune ipotesi di gaussianità
- un cenno ai metodi principali per giustificare l'ipotesi di gaussianità

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri
- l'analisi delle componenti principali (PCA) e una giustificazione tramite opportune variabili gaussiane
- qualche rudimento della regressione, in particolare il metodo dei minimi quadrati, giustificato tramite opportune ipotesi di gaussianità
- un cenno ai metodi principali per giustificare l'ipotesi di gaussianità
- un cenno al metodo di Laplace per approssimare densità generali con opportune gaussiane.

Il caso reale: definizione veloce

Si dice che una variabile aleatoria $X \in \mathbb{R}$ ha densità continua gaussiana se vale

$$p(X = x) \propto \exp\left(ax^2 + bx\right), \quad \text{per ogni } x \in \mathbb{R},$$

per degli opportuni parametri $a, b \in \mathbb{R}$.

- la densità è l'esponenziale di un polinomio di secondo grado dei possibili valori $x \in \mathbb{R}$.

Il caso reale: definizione veloce

Si dice che una variabile aleatoria $X \in \mathbb{R}$ ha densità continua gaussiana se vale

$$p(X = x) \propto \exp\left(\underset{0}{a}x^2 + \overset{\in \mathbb{R}}{b}x\right), \quad \text{per ogni } x \in \mathbb{R},$$

per degli opportuni parametri $a,\ b \in \mathbb{R}$.

- la densità è l'esponenziale di un polinomio di secondo grado dei possibili valori $x \in \mathbb{R}$.
- dovendo essere $\int_{-\infty}^{\infty} p(X = x) dx < \infty$, allora $a \in \mathbb{R}$ necessariamente deve essere $a < 0$

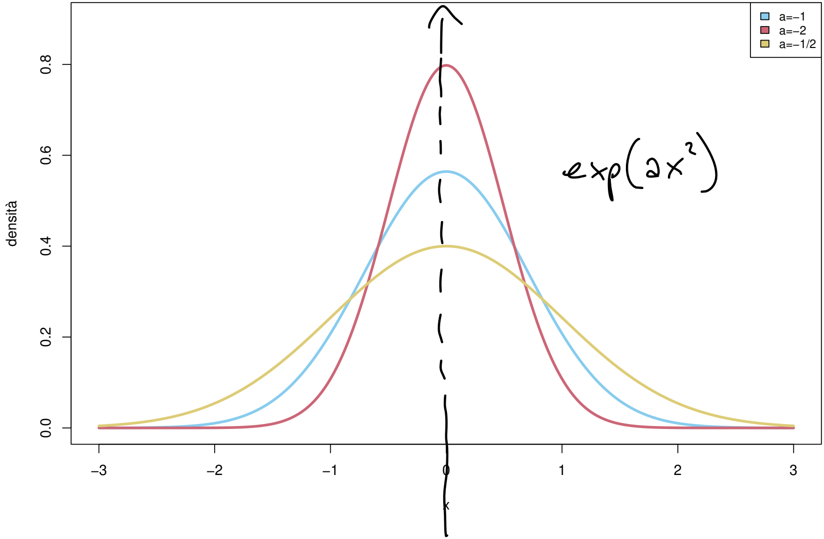


Figure 1: densità gaussiana al variare del parametro $a < 0$, $b = 0$

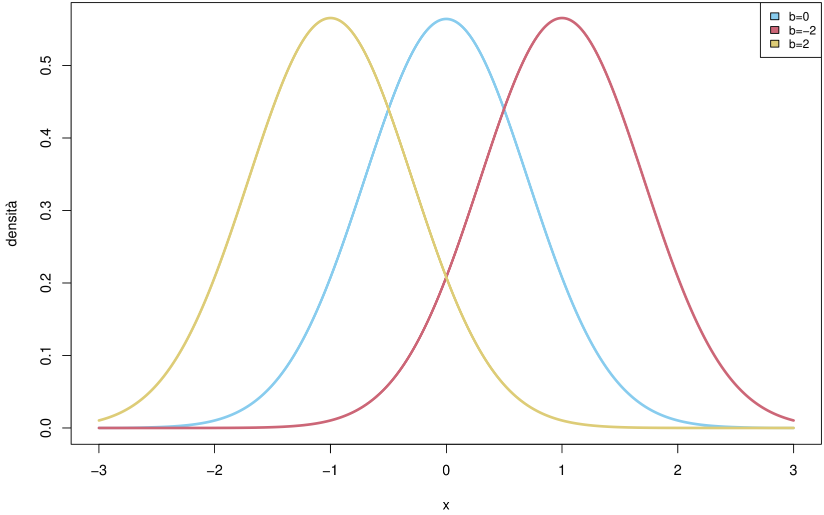


Figure 2: densità gaussiana al variare del parametro b, $a = 1$

Intepretazione dei parametri

Sia X una variabile con densità gaussiana

$$p(X = x) \propto \exp\left(ax^2 + bx\right).$$

Allora vale

$$a = -\frac{1}{2\sigma_X^2}, \quad b = \frac{\mathbb{E}[X]}{\sigma_X^2},$$

ossia

$$\boxed{\mathrm{Var}(X) = \sigma_X^2 = -\frac{1}{2a}} \quad \boxed{\mathbb{E}[X] = -\frac{b}{2a}.}$$

Dimostrazione $p(X=x) = \exp(+ax^2+bx)\cdot c$

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x\, p(X=x)\, dx = \int_{-\infty}^{+\infty} x \exp(ax^2+bx)\cdot c\, dx \qquad (1)$$

$$\boxed{\frac{d}{dx}\left(\exp(ax^2+bx)\right) = (2ax+b)\exp(ax^2+bx)}$$

$$\int_{-\infty}^{+\infty} \frac{d}{dx}\left(\exp(ax^2+bx)\cdot c\right) = \int_{-\infty}^{+\infty} (2ax+b)\exp(ax^2+bx)\cdot c$$

$$= 0 - 0 = 0$$

$$= 2a\int_{-\infty}^{+\infty} x\, p(X=x)\, dx + b\underbrace{\int_{-\infty}^{+\infty} p(X=x)\, dx}_{=1}$$

$$\boxed{2a\,\mathbb{E}[X] + b = 0}$$

$$\boxed{\mathbb{E}[X^2] = ?}$$

Definizione usuale

Si dice che $X \in \mathbb{R}$ ha densità continua gaussiana di valor medio $m \in \mathbb{R}$ e varianza $\sigma^2 > 0$, e si scrive brevemente $\mathcal{N}(m, \sigma^2)$, se

$$p(X = x) \propto \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right)$$

- Più esplicitamente, si può mostrare che vale l'identità

$$p(X = x) = \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right)\frac{1}{\sqrt{2\pi\sigma^2}}.$$

Definizione usuale

Si dice che $X \in \mathbb{R}$ ha densità continua gaussiana di valor medio $m \in \mathbb{R}$ e varianza $\sigma^2 > 0$, e si scrive brevemente $\mathcal{N}(m, \sigma^2)$, se

$$p(X = x) \propto \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right).$$

- Più esplicitamente, si può mostrare che vale l'identità

$$p(X = x) = \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right)\frac{1}{\sqrt{2\pi\sigma^2}}.$$

- La costante $1/\sqrt{2\pi}$ è interessante da calcolare analiticamente, ma non troppo utile nelle applicazioni.

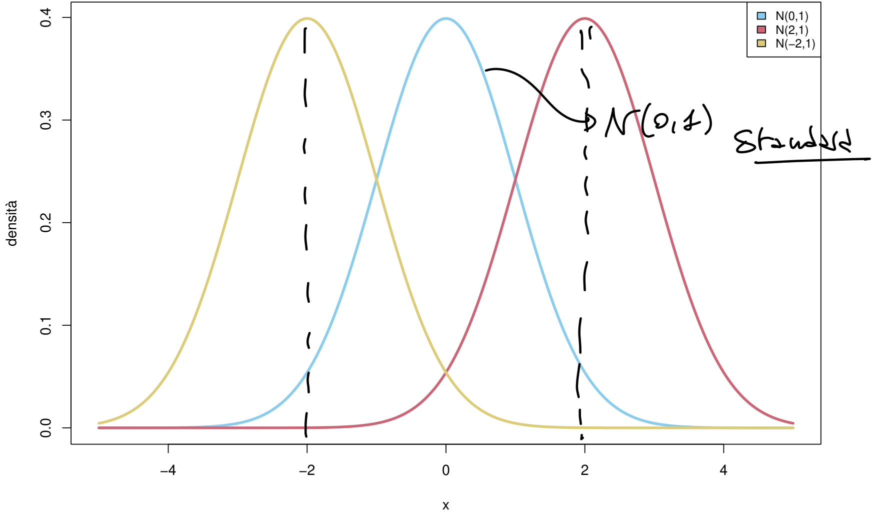


Figure 3: densità gaussiana al variare del parametro *m* (con $\sigma = 1$ costante

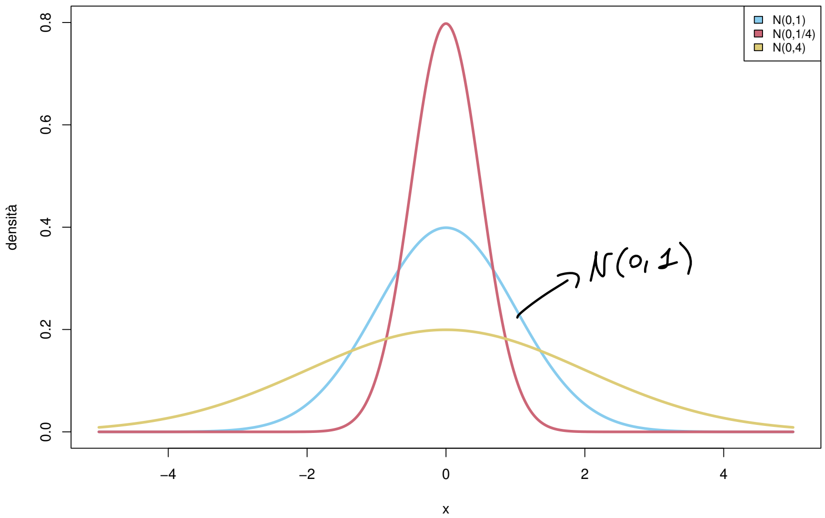


Figure 4: densità gaussiana al variare del parametro σ (con $m = 0$ costante

Proprietà di massima entropia

Al variare di tutte le possibili densità continue per una variabile X, $p(X = x)$, con $x \in \mathbb{R}$, tali che il valor medio e la varianza di X siano fissati

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xP(X = x)dx = m, \quad \mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-m)^2 P(X = x)dx = \sigma^2$$

la densità gaussiana $\mathcal{N}(m, \sigma^2)$ è quella di **massima entropia**.

- Pertanto, seguento principio di massima entropia, avendo a disposizione come informazione su una variabile aleatoria (reale) solamente il suo valor medio m e la varianza σ^2, sia imporrà che sia una densità gaussiana $\mathcal{N}(m, \sigma^2)$.

Trasformazione affine

Sia X una variabile con densità continua $\mathcal{N}(m, \sigma^2)$ e siano $\lambda \neq 0$, $c \in \mathbb{R}$. Allora la variabile $Y = \lambda X + c$ ha densità continua gaussiana, di parametri $\mathcal{N}(\lambda m + c, \lambda^2 \sigma^2)$.

- se X ha densità gaussiana $\mathcal{N}(m, \sigma^2)$, la sua standardizzata

$$\frac{X - m}{\sigma} \quad \text{ha densità continua } \mathcal{N}(0, 1),$$

pertanto detta anche densità **gaussiana standard**, che ha densità

$$\boxed{\exp\left(-\frac{1}{2}x^2\right) \frac{1}{\sqrt{2\pi}} \quad \text{per } x \in \mathbb{R}.}$$

Trasformazione affine

Sia X una variabile con densità continua $\mathcal{N}(m, \sigma^2)$ e siano $\lambda \neq 0$, $c \in \mathbb{R}$. Allora la variabile $Y = \lambda X + c$ ha densità continua gaussiana, di parametri $\mathcal{N}(\lambda m + c, \lambda^2 \sigma^2)$.

- se X ha densità gaussiana $\mathcal{N}(m, \sigma^2)$, la sua standardizzata

$$\frac{X - m}{\sigma} \quad \text{ha densità continua } \mathcal{N}(0, 1),$$

pertanto detta anche densità **gaussiana standard**, che ha densità

$$\exp\left(-\frac{1}{2}x^2\right) \frac{1}{\sqrt{2\pi}} \quad \text{per } x \in \mathbb{R}.$$

- Se $\lambda = 0$, la variabile $\lambda X + c = c$ è costante. Per uniformare le notazioni, si conviene di considerare anche le variabili costanti come caso *degenere* di una densità gaussiana.

Dimostrazione

$$X \sim N(m, \sigma^2) \qquad \begin{matrix}\lambda \neq 0 \\ c \in \mathbb{R}\end{matrix} \qquad Y = \lambda X + c$$

$$p(Y = y) = p(\lambda X + c = y) = p\left(X = \frac{y-c}{\lambda}\right)\frac{1}{|\lambda|}$$

$$= \exp\left(-\frac{1}{2}\frac{\left(\frac{y-c}{\lambda} - m\right)^2}{\sigma^2}\right)\frac{1}{\sqrt{2\pi\sigma^2}}\frac{1}{|\lambda|}$$

$$= \exp\left(-\frac{1}{2}\frac{(y - \overbrace{(c+\lambda m)}^{m'})^2}{\underbrace{\lambda^2\sigma^2}_{(\sigma')^2}}\right)\frac{1}{\sqrt{2\pi(\lambda\sigma)^2}}$$

CDF

$$CDF_{\mathcal{N}(m,\sigma^2)}(t) = \int_{-\infty}^{t} \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right)\frac{1}{\sqrt{2\pi\sigma^2}}\,dx$$

La funzione di ripartizione gaussiana (anche nel caso standard) non è esprimibile in termini di funzioni elementari.

- Il comando R per ottenerne i valori è `pnorm()`.

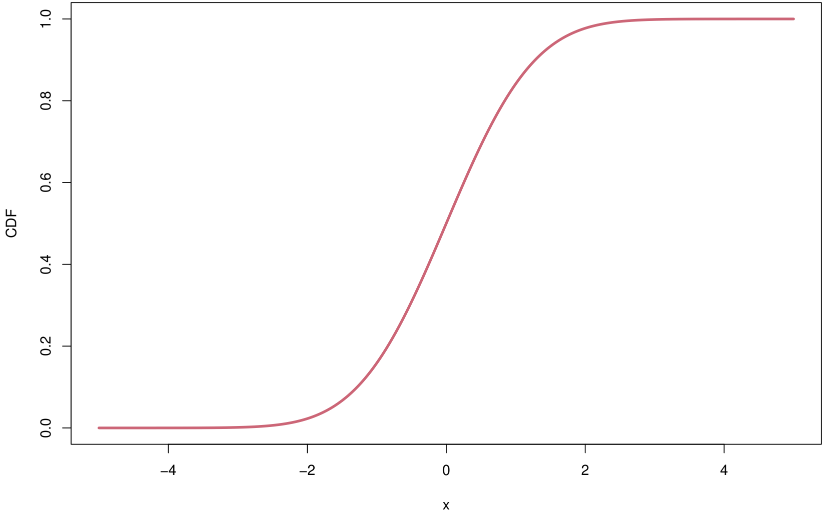


Figure 5: CDF di una variabile gaussiana standard

MGF e funzione caratteristica

Sia X una variabile con densità continua $\mathcal{N}(m, \sigma^2)$. Allora

$$\mathrm{MGF}_X(t) = \exp\left(mt + \frac{\sigma^2}{2}t^2\right),$$

e

$$\varphi_X(\xi) = \exp\left(im\xi - \frac{\sigma^2}{2}\xi^2\right).$$

$t \longleftrightarrow i\xi$

~~Dimostrazione~~ Applicazione

$$\mathbb{E}[X^2] = \frac{d^2}{(dt)^2} MGF(t)\Big|_{t=0}$$

$$MGF(t) = \exp\left(\frac{\sigma^2 t^2}{2} + mt\right) \qquad \frac{d}{dt} MGF(t) = \underline{MGF(t)\left(\sigma^2 t + m\right)}$$

$$\frac{d^2}{(dt)^2} MGF(t) = \frac{d}{dt}\Big(MGF(t)\,(\sigma^2 t + m)\Big) = MGF\,(\sigma^2 t + m)^2 + MGF\,\sigma^2$$

$\to t = 0$ $\qquad \mathbb{E}[X^2] = m^2 + \sigma^2 \qquad Var(X) = \mathbb{E}[X^2] - m^2 = \underline{\sigma^2}$ <u>ok</u>

Dimostrazione CASO $N(0,1)$ Standard

$$MGF(t) = \int_{-\infty}^{+\infty} e^{tx} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}} = \int_{-\infty}^{+\infty} \exp\left(tx - \frac{x^2}{2}\right) \frac{dx}{\sqrt{2\pi}}$$

$$= \int_{-\infty}^{+\infty} \exp\Big(-\frac{1}{2}\big(\underbrace{x^2 - 2tx + t^2}_{(x-t)^2} - t^2\big)\Big) \frac{dx}{\sqrt{2\pi}}$$

$$= \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}\frac{(x-t)^2}{1^2}\right) \exp\left(\frac{t^2}{2}\right) \frac{dx}{\sqrt{2\pi}}$$

$$= \underbrace{\int_{-\infty}^{+\infty} p\big(N(t,1) = x\big)\, dx}_{=1} \cdot \boxed{\exp\left(\frac{t^2}{2}\right)}$$

Dimostrazione

$$MGF_{aX+b}(t) = \mathbb{E}[\exp((aX+b)t)] = e^{bt}\, MGF_X(at)$$

$X \sim N(0,1)$ $\quad MGF_X(s) = \exp\left(\frac{s^2}{2}\right)$

$Y \sim N(m,\sigma^2)$ $\quad \hat{Y} = \frac{Y-m}{\sigma} \sim N(0,1)$

$$\boxed{Y = \sigma\hat{Y}+m}$$

$$MGF_{\sigma\hat{Y}+m}(t) = e^{tm}\exp\left(\frac{(\sigma t)^2}{2}\right) = \exp\left(\frac{\sigma^2 t^2}{2}+tm\right)$$

Problema 1

$\boxed{T = 8{:}25 \pm 5}$ $T \sim N(8{:}25, 25)$

T

L'orario d'arrivo a lezione degli studenti di ingegneria robotica segue approssimativamente una distribuzione gaussiana di media 8:25 e deviazione standard 5 minuti. Preso uno studente a caso,

1. calcolare la probabilità che arrivi dopo l'inizio delle lezioni (8:30);
2. calcolare il ritardo medio (in minuti).

Esprimere eventualmente i risultati come opportuni integrali o indicare un comando R per il calcolo numerico.

$P(\text{"arriva in ritardo"}) = P(\underline{T > 8{:}30})$

$\underline{T = 8{:}25 + 5\hat{T}}$ | $\hat{T} \sim N(0,1)$

$= P(\hat{T} > 1)$

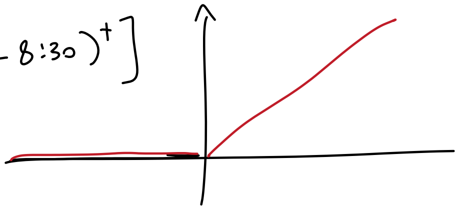
$= \int_1^{+\infty} \exp\left(-\frac{t^2}{2}\right) \frac{dt}{\sqrt{2\pi}}$

$= 1 - \text{pnorm}(1) \approx 16\%$

$R^* = T - 8{:}30 \qquad T \sim N(8{:}25, 5)$

$\mathbb{E}[R^*] = \mathbb{E}[T - 8{:}30] = 8{:}25 - 8{:}30 = -5 \text{ min}$

$R = (T - 8{:}30)^+ \qquad (z)^+ = \max\{0, z\}$



$\mathbb{E}[R] = \mathbb{E}[(8{:}25 + 5\hat{T} - 8{:}30)^+]$

$= \mathbb{E}[(5(\hat{T} - 1))^+]$

$= 5\,\mathbb{E}[(\hat{T} - 1)^+] = 5\int_{-\infty}^{+\infty} (t-1)^+ \exp\left(-\frac{t^2}{2}\right)\frac{dt}{\sqrt{2\pi}}$

$$5\int_{-\infty}^{+\infty}(t-1)^{+}\exp\left(-\frac{t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} = 5\int_{1}^{+\infty}(t-1)\exp\left(-\frac{t^2}{2}\right)\frac{dt}{\sqrt{2\pi}}$$

$$= 5\underbrace{\int_{1}^{+\infty} t\exp\left(-\frac{t^2}{2}\right)\frac{dt}{\sqrt{2\pi}}}_{-\frac{d}{dt}\exp\left(-\frac{t^2}{2}\right)} - 5\underbrace{\int_{1}^{+\infty}\exp\left(-\frac{t^2}{2}\right)\frac{dt}{\sqrt{2\pi}}}_{SUR_{\hat{T}}(1)}$$

$$= 5\left(-\exp\left(-\frac{t^2}{2}\right)\right)\Big|_{1}^{+\infty} - 5\,SUR_{\hat{T}}(1)$$

$$= 5\exp(-1/2) - 5\,SUR_{\hat{T}}(1)$$

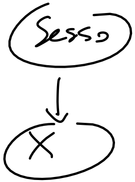
Problema 2

L'altezza degli studenti (maschi) del corso di ingegneria è rappresentata da una distribuzione gaussiana di media 175 cm e deviazione standard 10 cm. L'altezza delle studentesse (femmine) è pure una gaussiana di media 160 cm con deviazione standard 10 cm. Preso uno/a studente a caso, si osserva che è alto/a 165 cm. Dire se è più probabile che sia maschio o femmina,

1. senza conoscere la percentuale di studenti maschi e femmina nel corso;
2. sapendo anche che i maschi rappresentano il 70% degli studenti di ingegneria e il 30% è femmina (solo per semplicità di calcolo escludiamo le persone non identificate in uno dei due generi).

$$P(M) = P(F) = \frac{1}{2}$$

X altezza



$$P(M \mid X = 165) =$$

$$= \frac{P(M)\, p(X=165 \mid M)}{p(X=165)} \propto \exp\left(-\frac{1}{2}\left(\frac{165-175}{10}\right)^2\right)\frac{1}{\sqrt{\pi 10^2}}$$

$$P(F \mid X=165) = \frac{P(F)\, p(X=165 \mid F)}{p(X=165)} \propto \exp\left(-\frac{1}{2}\left(\frac{165-160}{10}\right)^2\right)\frac{1}{\sqrt{\ }}$$

$$\Longrightarrow \quad P(F \mid X=165) > P(M \mid X=165)$$

$P(F) = 30\%$

$P(M) = 70\%$

$$P(F \mid x=165) = 30\% \exp\left(-\frac{1}{2}\left(\frac{165-160}{10}\right)^2\right) = \frac{3}{10}\exp\left(-\frac{1}{8}\right)$$

$$P(M \mid x=165) = 70\% \exp\left(-\frac{1}{2}\left(\frac{165-175}{10}\right)^2\right) = \frac{7}{10}\exp\left(-\frac{1}{2}\right)$$