1 Descrizione sommaria degli argomenti svolti a lezione, anno 2006

Lezione 1 (21/2, 2 ore). Introduzione al corso (testo: S. Ross, materiale complementare in rete alle pagine di Gubinelli e Flandoli). Si comincia dal Capitolo 3: oggetti del calcolo delle probabilità: esiti (eventi elementari), spazio degli esiti, eventi ed operazioni su di essi, probabilità e sue regole. Probabilità condizionale, formula di fattorizzazione, interpretazione con albero, esercizio 1.1 del 22/7/03.

Lezione 2 (22/2, 2 ore). Formula di Bayes. Soluzione dell'es. 5 del 17/5/03, con domanda preliminare: calcolare la probabilità che quel prodotto venga acquistato. Uso dell'albero in relazione alle probabilità delle cause. Assegnati per casa gli esercizi 1 (i, ii) del 4/6/02, ed il 29 del Ross.

Spazi equiprobabili, principio di enumerazione, esercizio n. 19 del Ross. Permutazioni (applicazioni biunivoche), loro cardinalità $= n!$. Stringhe di 0-1 lunghe n: la loro cardinalità è 2^n; il numero di quelle con k elementi pari ad uno è $\binom{n}{k}$, coefficiente binomiale.

Lezione 3 (23/2, 1 ora). Capire che il numero di insiemi di k elementi tratti da un insieme di n elementi è $\binom{n}{k}$.

Definizione di indipendenza tra due eventi A e B, prima a partire dalla probabilità condizionale ($P(A) = P(A|B)$), poi nella forma simmetrica $P(A \cap B) = P(A)P(B)$. Esercizio 1.3 del 22/7/03.

Lezione 4 (28/2, 2 ore). Esercitazione. Vengono risolti e commentati i punti i) e ii) dell'es. 1 del compito del 4/6/02 sia con l'approccio dell'albero sia tramite formula di fattorizzazione e Bayes. Esercizi 18, 20, 23, 24 e 27 del Ross.

Lezione 5 (1/3, 2 ore). Esercitazione. Risolti gli esercizi 6 e 9 della lista e l'esercizio 37, prima domanda, del Ross. Osservazioni sulle probabilità condizionate: si comportano come probabilità. Consigliati per casa gli esercizi 1-9 della lista, 25, 29, 31, 35, 37 (seconda e terza parte) e 41 del Ross (risolti in rete) ed eventualmente anche 34, 36, 38 del Ross. Inoltre quasi tutti i primi esercizi dei compiti sono già risolubili, ad es. del 1/9/02.

Inizio dello studio dei capitoli 4 e 5 del Ross. Variabili aleatorie discrete. Alcuni esempi di variabili aleatorie che si incontrano nella pratica (a parole). Variabili aleatorie discrete: loro descrizione tramite i valori e la massa (o distribuzione) di probabilità. Rappresentazione grafica della massa di probabilità e varie notazioni per descrivere una variabile discreta. Esempi delle v.a.

di Bernoulli e binomiale (verifica della somma pari ad uno). Loro grafico.

Lezione 6 (2/3, 1 ora). Concetto di indipendenza di variabili aleatorie (nel caso discreto), come sottoprodotto del concetto di indipendenza di eventi. Criterio per la verifica dell'indipendenza (due v.a. discrete sono indipendenti se la probabilità che entrambi prendano determinati valori è uguale al prodotto delle probabilità per le singole variabili). Trasformazione di v.a. discrete (osservazioni nel caso di non iniettività). Se X ed Y sono indipendenti ed f, g sono funzioni, allora $f(X)$ e $g(Y)$ sono indipendenti.

Definizione di funzione di distribuzione cumulativa $F(x)$ e suo grafico per una v.a. discreta (con un numero finito di valori).

Lezione 7 (7/3, 2 ore). Concetto di variabile aleatoria come funzione definita sullo spazio degli esiti, che ad ogni esito associa un numero; e sulla versione intuitiva della definizione, come grandezza aleatoria che assume certi valori con certe probabilità. Richiamo su v.a. di Bernoulli, $X \sim B(1,p)$, su v.a. indipendenti, sui coefficienti binomiali (osservando alcuni fatti, ad esempio $\binom{n}{k} = \binom{n}{n-k}$, $0! = 1$, $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{1} = \binom{n}{n-1} = n$), sulle v.a. binomiali, $X \sim B(n,p)$: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, ..., n$. Teorema: date $X_1, ..., X_n \sim B(1,p)$ indipendenti, posto $S = X_1 + ... + X_n$, S risulta essere una $B(n,p)$. Dimostrazione, calcolando $P(S = k)$ come somma degli esiti con k successi, che sono $\binom{n}{k}$, ciascuno con probabilità $p^k (1-p)^{n-k}$.

Esempio di una banca con 1000 correntisti, indipendenti, ciascuno con probabilità 1/5 di presentarsi in un generico giorno; si deve calcolare la probabilità di avere più di 300 visite in un giorno. Altri esercizi.

Lezione 8 (8/3, 2 ore) Teorema degli eventi rari: se $X \sim \mathrm{Bin}(n,p)$ con $n \to \infty$, $np = \lambda$ allora $P(X = k) \to e^{-\lambda} \lambda^k / k!$, $k = 0, 1, \ldots$. Definizione di v.a. di Poisson di parametro λ: $P(N = k) = e^{-\lambda} \lambda^k / k!$, $k = 0, 1, \ldots$. Teorema degli eventi rari come convergenza della distribuzione di v.a. Binomiali alla distribuzione di v.a. di Poisson. Approssimazione di Binomiali con Poisson.

Definizione di valor medio di una v.a. discreta; sua interpretazione grafica; esempi elementari; sua motivazione in relazione alla media aritmetica di una campione sperimentale ($\frac{1}{n} \sum x_i = a_1 \frac{n_1}{n} + ... \sim \sum a_k p_k$).

Lezione 9 (9/3, 1 ora). Valor medio di una trasformazione $f(X)$. Caso particolare: definizione di Varianza e sua interpretazione. Definizione di deviazione standard e sua interpretazione anche grafica. Simboli μ, σ^2, σ. Proprietà del valor medio: linearità (senza dimostrazione).

Lezione 10 (14/3, 2 ore). $Var[X] = E[X^2] - \mu^2$, con dimostrazione.

Esempio di calcolo della varianza di una Bernoulli nei due modi. Teorema: X, Y indipendenti implica $E[XY] = E[X]E[Y]$. Corollario: X, Y indipendenti implica $Var[X+Y] = Var[X] + Var[Y]$. In generale, introdotta la Covarianza, vale $Var[X+Y] = Var[X] + Var[Y] + 2Cov(X,Y)$. Esercizio: calcolo di media e varianza di una binomiale, usando la rappresentazione come somma di Bernoulli indipendenti. Calcolo (un po' intuitivo) di media e varianza di una Poisson.

Dimostrazione del teorema. Funzione di massa congiunta e marginali. Relazione generale $p_X(a) = \sum_b p(a,b)$ ecc., relazione particolare $p(a,b) = p_X(a)\, p_Y(b)$ per v.a. indipendenti. Formula generale $E[g(X,Y)] = \sum_{a,b} g(a,b)\, p(a,b)$. Da questo volendo si potrebbe dimostrare la linearità.

Lezione 11 (15/3, 2 ore) Esercizio del compitino del 20/4/04.

Funzione generatrice dei momenti (FGM) di una v.a. Definizione $\phi_X(t) = E[e^{tX}]$ e calcolo per le v.a. di Bernoulli. Proprietà principali: le derivate in zero della $\phi_X(t)$ danno il valor medio delle potenze di X: $\phi'(0) = E[X]$, $\phi''(0) = E[X^2]$; la FGM della somma di due v.a. indipendenti è uguale al prodotto delle rispettive FGM. Dimostrazione di queste proprietà. Calcolo della FGM per la Binomiale usando il fatto che la Binomiale è somma di Bernoulli.

Lezione 12 (16/3, 1 ora) Calcolo diretto della FGM per la Poisson e calcolo di media e varianza. Esercizi per casa: 1) verificare media e varianza di una binomiale usando la FGM; 2) compito del 20/4/04 sulla FGM; 3) verificare che $\phi(0) = 1$ ed usarlo per stabilire che certe funzioni non possono essere FGM; 4) verificare che la generatrice di una $B(n,p)$ tende a quella di una $\mathcal{P}(\lambda)$ nel regime del teorema degli eventi rari; 5) definizione di classe autoriproducente e verifica che le Bernoulli non lo sono, mentre le binomiali e le Poisson sì.

Lezione 13 (21/3, 2 ore) Proprietà di concentrazione delle $B(n,p)$ con n grande (p non piccolo): σ è molto minore di n e di μ. Spiegazione dell'esercizio della banca.

Normalizzazione di una v.a.: se X ha valor medio μ e varianza σ^2 allora la v.a. $Z = (X-\mu)/\sigma$ ha media nulla e varianza unitaria. Dimostrazione.

Esercizio 2 del compito del 4/6/02 ed esercizio 2 del compito del 22/9/04, entrambe sulle v.a. discrete.

Viene suggerito il seguente problema: abbiamo una centrale telefonica con 1000 clienti. In un tipico minuto c'è probabilità pari a 1/50 che un utente richieda l'utilizzo di una linea per una telefonata. Quante linee deve avere la centrale per garantire il servizio con probabilità pari a 0.99?

Si inizia lo studio delle v.a. continue. Una v.a. X è (assolutamente) continua se la sua funzione di ripartizione F_X si scrive come l'integrale di una funzione $f_X(x) \geq 0$: $F_X(t) = \int_{-\infty}^{t} f_X(x)dx$. La funzione $f_X(x)$ si chiama densità di probabilità e soddisfa: $\int_{-\infty}^{\infty} f_X(x)dx = 1$. Definizione di valor medio per una v.a. continua $E[X] = \int x f_X(x)dx$ e inoltre il valor medio di una funzione generica di una v.a. continua: $E[g(X)] = \int g(x)f_X(x)dx$. Osservazioni sull'analogia tra operazioni sulle v.a. continue e le corrispondenti operazioni sulle v.a. discrete. Osservazione sul fatto che la funzione di ripartizione di una v.a. continua è una funzione continua e inoltre $f_X(t) = dF_X(t)/dt$ quando F è differenziabile.

Definizione di v.a. Uniforme. X è una v.a. uniforme in $[\alpha, \beta]$ se la sua funzione di densità è nulla fuori $[\alpha, \beta]$ ed è costante all'interno. Valor medio e varianza di una v.a. uniforme e calcolo della funzione generatrice dei momenti.

Lezione 14 (22/3, 2 ore) Esercizi sulle v.a. continue definite tramite la densità: trovare la funzione di ripartizione, calcolare il valor medio. Esempio di come procedere per calcolare la densità di una v.a. definita in funzione di un'altra v.a. continua di cui si conosce la densità (vedere p. 117 del Ross). Si parte da X uniforme in $[0, 1]$ e si vuole trovare la densità di probabilità di $Y = X^2$. Passaggi chiave: determinare la F_Y e poi derivarla per ottenere f_Y.

Variabili aleatorie esponenziali. Formula $\int_a^b e^{-\lambda x}dx = \frac{e^{-\lambda a} - e^{-\lambda b}}{\lambda}$. Determinazione della costante nella formula della densità. Formula $P(X > t) = e^{-\lambda t}$. Funzione di distribuzione cumulativa.

Lezione 15 (23/3, 1 ora). Teorema sulla proprietà di mancanza di memoria: enunciato, interpretazione dell'enunciato (pensando all'istante aleatorio in cui acade un certo evento), dimostrazione. Calcolo della media, $\frac{1}{\lambda}$ (vale quindi $\lambda = \frac{1}{E[X]}$) e della varianza, $\frac{1}{\lambda^2}$ (quindi la deviazione σ è pari alla media). Calcolo della generatrice, $\phi(t) = \frac{\lambda}{\lambda - t}$, definita solo per $t < \lambda$.

Lezione 16 (28/3, 2 ore). Funzione generatrice delle v.a. esponenziali. Proprietà di mancanza di memoria delle v.a. esponenziali, enunciato, dimostrazione e interpretazione per i tempi di vita di sistemi senza usura.

Soluzione dell'esercizio sulle v.a. autoriproducenti, per Bernoulli, binomiale, Poisson, esponenziale.

Lezione 17 (29/3, 2 ore). Esercitazione in aula (primo compitino). Variabili aleatorie gaussiane: definizione nel caso canonico e nel caso generale, grafico, ruolo di μ e σ^2 a livello grafico, calcolo delle sua funzione generatrice di una $X \sim N(\mu, \sigma^2)$: $\varphi(t) = e^{t\mu + \frac{t^2\sigma^2}{2}}$

Lezione 18 (30/3, 1 ora). Gaussiane: cenno al calcolo di media e varianza senza generatrice; funzione di distribuzione cumulativa $F(x)$, notazione $\Phi(x)$ nel caso canonico, non calcolabilità analitica ed uso delle tavole, $\Phi(x) = 1 - \Phi(-x)$, $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ nel caso generale.

Teorema: X, Y gaussiane indipendenti, α, β, γ numeri reali, implica $\alpha X + \beta Y + \gamma$ gaussiana. Si dimostra usando la generatrice ed i suoi teoremi: quello noto sulla generatrice della somma di indipendenti, e quello che dice $\varphi_{\alpha X+\beta}(t) = e^{\beta t}\varphi_X(\alpha t)$, dimostrati. Le gaussiane sono quindi autoriproducenti. Media e varianza di $\alpha X + \beta Y + \gamma$ si trova dal calcolo appena svolto sulla generatrice $\varphi_{\alpha X+\beta Y+\gamma}(t)$ ma anche più facilmente dalle proprietà note su media e varianza.

Lezione 19 (4/4, 2 ore). Concetto di quantile, inverso della distribuzione cumulativa. Quantile gaussiano canonico di ordine α (notazione q_α), sua lettura tramite le tavole, formula di legame col quantile di una gaussiana generica, $q_{1-\alpha} = -q_\alpha$, $\Phi(q_\alpha) = \alpha$, $q_{\Phi(x)} = x$; notazione z_α: quel numero tale che $P(Z > z_\alpha) = \alpha$, con $Z \sim N(0,1)$; vale $z_\alpha = q_{1-\alpha}$. Esercizi vari di calcolo di probabilità sulle gaussiane, in cui ci si riconduce a Φ o quantili.

Teorema limite centrale: enunciato, esercizi di applicazione.

Lezione 20 (5/4, 2 ore) Esercizio su v.a. discrete: data X che vale $0,1,2$ con ugual probabilità, Y che vale $0,1$ con ugual probabilità, indipendenti, 1) calcolare μ_X, σ_X^2, $\phi_X(t)$; 2) calcolare $E[XY - Y^2 + 2Y]$; 3) calcolare $P(X+Y) = 1$.

Primi elementi di statistica: il concetto di campione $X_1, ..., X_n$; se $X_1, ..., X_n$ è un campione di numerosità n estratto da una v.a. X avente media μ e varianza σ^2, chiamiamo media aritmetica la v.a. $\overline{X}_n = \frac{X_1+...+X_n}{n}$. Calcolo con le proprietà dei valori medi: $\overline{X}_n$ ha media μ e varianza $\frac{\sigma^2}{n}$. Se X è gaussiana, lo è anche $\overline{X}_n$. Interpretazione del fatto che $\overline{X}_n$ è concentrata intorno a μ, cioè che $Var\left[\overline{X}\right] = \frac{\sigma^2}{n}$ è molto minore di σ^2: i valori sperimentali di $\overline{X}_n$ sono generalmente più vicini a μ (quindi sono una miglior stima di μ), di quanto non siano i valori di X stessa.

Esercizi svolti in questo periodo: 1) calcolare $q_{0.9842}$, $q_{0.995}$, $\Phi(-2.15)$; 2) per una $X \sim N(5;16)$ trovare x tale che $P(X \leq x) = 0.8$; in generale, per una $X \sim N(\mu, \sigma^2)$, dato $\alpha \in (0,1)$, trovare λ tale che $P(X \leq \lambda) = \alpha$; 3) se $X \sim N(1;1)$, calcolare $P(X > 3)$; se $X \sim N(5;4)$, calcolare $P(|X-5| > 1)$; 4) se $X \sim N(7;8)$, trovare λ tale che $P(X > \lambda) = 0.1$; se $X \sim N(11;25)$, trovare λ tale che $P(|X-11| < \lambda) = 0.1$; 5) esercizio 27 della lista (sul teorema limite centrale).

Si suggerisce inoltre di corredare la tabella dei quantili del libro con i seguenti: $q_{0.9} = 1.282$, $q_{0.95} = 1.645$, $q_{0.975} = 1.960$, $q_{0.99} = 2.326$, $q_{0.995} = 2.576$, $q_{0.999} = 3.090$, $q_{0.9995} = 3.291$, $q_{0.99995} = 3.891$.

Lezione 21 (6/4, 1 ora) Indirizziamoci ad una valutazione quantitativa dell'errore che si commette approssimando μ con $\overline{X}$. Calcolo di $P\left(\left|\overline{X}_n \to \mu\right| < \varepsilon\right)$, nel caso gaussiano, a titolo di esercizio, che verrà ripreso.

20/4, Legge dei grandi numeri: enunciato sia con la convergenza in media quadratica: $E\left[\left(\overline{X}_n \to \mu\right)^2\right] \to 0$, sia con la convergenza in probabilità: $\forall \varepsilon > 0$, $P\left(\left|\overline{X}_n \to \mu\right| > \varepsilon\right) \to 0$. Dimostrazione in media quadratica in base al fatto che $\overline{X}_n$ ha media μ e varianza $\frac{\sigma^2}{n}$. Dimostrazione in probabilità usando i lemmi.

Lezione 22 (26/4, 2 ore). Esercizio su una soglia: se la domanda X è una gaussiana $N\left(\mu, \sigma^2\right)$, quanto bisogna tenere in magazzino per soddisfarla con probabilità 95%. Prima nel caso $X \sim N(0,1)$: detta θ tale scorta di sicurezza, deve valere $P(X \leq \theta) = 0.95$, quindi per definizione di quantile vale $\theta = q_{0.95}$. Nel caso generale si imposta l'equazione $P(X \leq \theta) = 0.95$, che equivale a $P\left(\frac{X-\mu}{\sigma} \leq \frac{\theta-\mu}{\sigma}\right) = 0.95$ dove ora $Z = \frac{X-\mu}{\sigma}$ è $N(0,1)$; quindi $\frac{\theta-\mu}{\sigma} = q_{0.95}$, ovvero $\theta = \mu + \sigma q_{0.95}$.

Terminologia: $\overline{X}$ è uno stimatore corretto (non distorto) di μ: $E\left[\overline{X}\right] = \mu$. Stimatori di σ^2:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2, \quad S_\mu^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2$$

Dimostrazione che entrambi sono non distorti (di conseguenza $\frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$ sarebbe distorto).

Valutazione quantitativa dell'errore che si commette approssimando μ con $\overline{X}$: una stima *certa* del tipo $\left|\overline{X} - \mu\right| \leq \delta$ è impossibile; cerchiamo di stabilire un'afferemazione del tipo $\left|\overline{X} - \mu\right| \leq \delta$ con probabilità $1 - \alpha$. Esercizio preliminare: data $X \sim N(0,1)$, dato $\alpha \in (0,1)$, trovare δ tale che valga $|X| \leq \delta$ con probabilità $1 - \alpha$. Soluzione: $\delta = q_{1-a/2}$.

Risolviamo il problema precedente nel caso più generale $X \sim N\left(\mu, \sigma^2\right)$ (rimandiamo il caso di una v.a. X ancor più generale, che tratteremo in modo approssimato col teorema limite centrale). Ricordiamo che le combinazioni lineari di gaussiane indipendenti sono gaussiane, quindi $\overline{X}$ è gaussiana. Inoltre, ricordiamo che $E\left[\overline{X}\right] = \mu$, $Var\left[\overline{X}\right] = \frac{\sigma^2}{n}$, quindi $Z = \sqrt{n}\frac{\overline{X}-\mu}{\sigma} \sim N(0,1)$. Pertanto $\left|\sqrt{n}\frac{\overline{X}-\mu}{\sigma}\right| \leq q_{1-a/2}$ con probabilità $1 - \alpha$.

Ovvero: $\left|\overline{X}-\mu\right| \leq \delta$ con probabilità $1-\alpha$, dove

$$\delta = \frac{\sigma q_{1-a/2}}{\sqrt{n}}.$$

Questa si dice *precisione*, mentre $1-\alpha$ si dice *confidenza*; inoltre l'intervallo $\left[\overline{X}-\delta, \overline{X}+\delta\right]$ in cui si trova μ con probabilità $1-\alpha$ si dice *intervallo di confidenza.* Scriveremo anche

$$\mu = \overline{X} \pm \delta \text{ con confidenza } 1-\alpha.$$

Lezione 23 (27/4, 1 ora). Completamento della teoria dell'intervallo di confidenza ed esempi.

Lezione 24 (2/5, 2 ore). Esercizi vari di preparazione al compitino (calcolo di quantili, probabilità gaussiane, teorema limite centrale). In particolare, confronto quantitativo tra binomiale, approssimazione con Poisson ed approssimazione gaussiana:

$$\sum_{k=0}^{10} \binom{100}{k} 0.05^k 0.95^{100-k} = 0.98853, \quad e^{-5} \sum_{k=0}^{10} \frac{5^k}{k!} = 0.9863,$$
$$\frac{10-5}{10\sqrt{0.05 \cdot 0.95}} = 2.2942 \to 0.989$$

$$\sum_{k=0}^{2} \binom{100}{k} 0.01^k 0.99^{100-k} = 0.92063, \quad e^{-1} \sum_{k=0}^{2} \frac{1^k}{k!} = 0.9197,$$
$$\frac{2-1}{10\sqrt{0.01 \cdot 0.99}} = 1.005 \to 0.8413$$

$$\sum_{k=0}^{20} \binom{100}{k} 0.3^k 0.7^{100-k} = 0.0164, \quad e^{-30} \sum_{k=0}^{20} \frac{30^k}{k!} = 0.035285,$$
$$\frac{20-30}{10\sqrt{0.3 \cdot 0.7}} = -2.1822 \to 0.015.$$

Commenti relativi su cosa rende peggiore l'approssimazione.

Lezione 25 (3/5, 2 ore). Esercizi (II compitino).

Lezione 26 (4/5, 1 ora). Calcolo della numerosità campionaria necessaria per avere una certa precisione, esempio di DOE (Design Of Experiments). Osservazione sul fatto che la precisione aumenta relativamente poco all'aumentare della numerosità.

Problema di valutare se un certo fattore ha influenza su un prodotto: ci indirizziamo al concetto di test statistico, che verrà spiegato successivamente.

Lezione 27 (9/5, 2 ore). Esercizi.

Lezione 28 (10/5, 2 ore). Test (o verifica) di ipotesi. Ipotesi nulla ed ipotesi alternativa. Regione critica (o di rigetto). Significatività. Logica alla base del test. Uso di $\left|\overline{X}_n - \mu_0\right|$ ed uso di $|z|$ per eseguire il test. Procedura di test. Esercizi.

Valore p: regione dei valori di α per cui il test risulta significativo, valore critico.

Lezione 29 (11/5, 1 ora). Errore di prima specie ed errore di seconda specie. Probabilità dell'errore di prima specie: si trova α. Specifica della condizione H_2 per il calcolo della probabilità di errore di seconda specie. Calcolo e formula finale. Potenza di un test.

Lezione 30 (16/5, 2 ore). Esercizi sui test. Varianti unilaterali.

Lezione 31 (17/5, 2 ore). III compitino.

Lezione 32 (23/5, 2 ore). Richiami e complementi sulla covarianza: linearità rispetto ai due argomenti; indipendenza implica covarianza nulla; il viceversa non vale in generale ma vale almeno per coppie (X, Y) gaussiane; definizione di covarianza sperimentale, calcolata a partire da un campione sperimentale $(x_1, y_1), \ldots, (x_n, y_n)$; difficoltà di utilizzazione delle informazioni quantitive sulla covarianza sperimentale.

Definizione di coefficiente di correlazione $\rho(X, Y)$. Vale $\rho(X, Y) \in [-1, 1]$. Non dipende dall'unità di misura: $\rho(\lambda X, \lambda Y) = \rho(X, Y)$. Il legame con l'indipendenza è lo stesso della covarianza. $Y = \alpha X + \beta + \varepsilon$, ε v.a. (che rappresenta l'errore) a media nulla e varianza σ^2_{err}, indipendente da X: vale $\rho(X, Y) = \frac{\alpha\sigma^2_X}{\sigma_X\sqrt{\alpha^2\sigma^2_X + \sigma^2_{err}}}$. Se $\sigma^2_{err} = 0$, $\rho(X, Y) =$ segno di α (1 se $\alpha > 0$, -1 se $\alpha < 0$). Se σ^2_{err} aumenta, $|\rho(X, Y)|$ diminuisce ($\rho(X, Y)$ si avvicina a zero). Indichiamo con $\widehat{\rho}(X, Y)$ il coefficiente di correlazione sperimentale. Tutte queste cose giustificano il criterio empirico: $\widehat{\rho}(X, Y)$ vicino a zero: campione sparpagliato nel piano; $\widehat{\rho}(X, Y)$ vicino ad 1: campione abbastanza concentrato lungo una retta avente coefficiente angolare positivo; $\widehat{\rho}(X, Y)$ vicino ad -1: campione abbastanza concentrato lungo una retta avente coefficiente angolare negativo.

Regressione. Dato un campione sperimentale (x_1, y_1), ... , (x_n, y_n), si cerca una legge funzionale $Y = f(X, \varepsilon)$ che lo descriva bene. Ipotizzando che tale legge funzionale sia lineare, della forma $Y = \alpha X + \beta + \varepsilon$, si trovano in modo approssimato i coefficienti α, β, e volendo anche σ^2_{err}, a partire dai dati sperimentali. Infatti deve valere $Cov\,(X, Y) = \alpha\sigma^2_X$, da cui si trova $\alpha = \frac{Cov(X,Y)}{\sigma^2_X}$; inoltre vale $\mu_Y = \alpha\mu_X + \beta$, da cui si trova $\beta = \mu_Y - \alpha\mu_X$. Da queste si decidono le regole sperimentali $\widehat{\alpha} = \frac{\widehat{Cov}(X,Y)}{S^2_X}$, $\widehat{\beta} = \overline{y} - \widehat{\alpha}\overline{x}$. Infine, da $\sigma^2_Y = \alpha^2\sigma^2_X + \sigma^2_{err}$ si trova $\sigma^2_{err} = \sigma^2_Y - \alpha^2\sigma^2_X$ e si decide la stima $S^2_{err} = S^2_Y - \widehat{\alpha}^2 S^2_X$.

Lezione 33 (24/5, 2 ore). Si riprende il discorso generale della ricerca di un legame funzionale $Y = f(X, \varepsilon)$ che descriva dei punti dati (x_1, y_1), ... , (x_n, y_n). Bisogna specificare la classe di funzioni (es. lineari, polinomiali di grado ≤ 3, esponenziali, ecc.) in cui cercare tale legame. Bisogna poi scegliere un concetto di distanza tra i punti sperimentali ed una funzione data, e cercare la funzione, della classe scelta, che rende minima tale distanza. Se si fa questo nel caso lineare con un'opportuna distanza, si ritrovano le formule per $\widehat{\alpha}$ e $\widehat{\beta}$ scritte sopra.

La teoria lineare basata sul modello $Y = \alpha X + \beta + \varepsilon$ si utilizza anche per studiare legami non lineari, ad esempio del tipo $Y = \alpha X^2 + \beta + \varepsilon$. Basta utilizzare i dati (x_1^2, y_1), ... , (x_n^2, y_n). Di fronte ad un insieme di dati (x_1, y_1), ... , (x_n, y_n), prima di tutto conviene raffigurarli in modo da intuire i possibili legami funzionali ragionevoli. Poi se ad esempio sono ragionevoli sia il legame lineare sia quello parabolico del tipo $Y = \alpha X^2 + \beta + \varepsilon$, si calcola il coefficiente di correlazione dei dati (x_1, y_1), ... , (x_n, y_n) e quello dei dati (x_1^2, y_1), ... , (x_n^2, y_n) e si sceglie il modello col coefficente migliore. Poi si calcolano i coefficienti del modello con le formule viste sopra. Per legami più complessi, con più coefficienti, serve la teoria della regressione multipla.

Carte di controllo della qualità: esempio che si trova in rete.