

**Normal distribution**  $\xi = N(\mu, \sigma^2) : f_\xi(x) = e^{-(x-\mu)^2/(2\sigma^2)} / (\sqrt{2\pi}\sigma)$ ,

$$E[\xi] = \mu \quad Var(\xi) = \sigma^2, \quad \varphi_\xi(t) = \exp(i\mu t - \sigma^2 t^2/2), \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} dy$$

$$e^{-x^2/2}/[(x+1/x)\sqrt{2\pi}] \leq 1 - \Phi(x) \leq e^{-x^2/2}/(x\sqrt{2\pi}), \quad x > 0$$

**Def.** An  $R^n$ -valued random variable  $\xi = (\xi_1, \dots, \xi_n)$  is Gaussian (or Multivariate Normal) if every linear combination  $\alpha_1 \xi_1 + \dots + \alpha_n \xi_n$  has a one dimensional Normal distribution.

$\xi$  is an  $R^n$ -valued Gaussian random variable if and only if its characteristic function has the form  $\varphi_\xi(t) = \exp(i \langle t, \mu \rangle - \frac{1}{2} \langle t, Q t \rangle)$  where  $Q$  is an  $n \times n$  symmetric nonnegative semi-definite matrix ( $Q$  is then covariance matrix of  $\xi$  and  $\mu$  is the mean of  $\xi$ ).

$$Cov(\xi, \eta) = E[(\xi - E[\xi])(\eta - E[\eta])] \quad \rho(\xi, \eta) = \frac{Cov(\xi, \eta)}{\sqrt{Var(\xi)Var(\eta)}}$$

**Characteristics function**  $\varphi_{\xi_1, \dots, \xi_n}(t_1, \dots, t_n) = E[e^{i(t_1 \xi_1 + \dots + t_n \xi_n)}] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i(t_1 x_1 + \dots + t_n x_n)} dP_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$

**Student's t-distribution**  $t_n = \frac{\xi}{\sqrt{\chi_n^2/n}}$ , where  $\xi \sim N(0,1)$ ,  $\chi_n^2$  -chi-square distribution with  $n$  degrees of freedom,  $\xi \perp \chi_n^2$ ,  $\chi_n^2 = \xi_1^2 + \dots + \xi_n^2$ ,  $\xi_1 \perp \xi_2 \perp \dots \perp \xi_n$ ,  $\xi_i \sim N(0,1)$   $i = 1, \dots, n$

**Chebyshev's inequality**  $P\{|\xi - \mu| \geq a\} \leq Var(\xi)/a^2$

**Convergence**  $\xi_n \xrightarrow{P} \xi : P(|\xi_n - \xi| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0 \quad \forall \varepsilon > 0$

**(almost sure)**  $\xi_n \xrightarrow{a.s.} \xi : P(\omega \in \Omega | \xi_n(\omega) \rightarrow \xi(\omega)) = 1$

**Def.** A sequence  $\xi_1, \dots, \xi_n, \dots$  of random variables is said to **converge in distribution**, or **converge weakly**, or **converge in law** to a random variable  $\xi$  ( $\xi_n \xrightarrow{w-d} \xi$ )

if  $F_{\xi_n}(x) \xrightarrow{n \rightarrow \infty} F_\xi(x)$  for every number  $x \in R$  at which  $F_\xi(x)$  is continuous.  $\xi_n \xrightarrow{w} \xi$  if and only if  $E[g(\xi_n)] \xrightarrow{n} E[g(\xi)]$  for all continuous bounded functions  $g$ .

**Central Limit Theorem (CLT)**

$$P\left\{\frac{\xi_1 + \dots + \xi_n - n\mu}{\sigma\sqrt{n}} \leq x\right\} \rightarrow \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} dy \quad n \rightarrow \infty, \text{ where}$$

$\xi_1, \dots, \xi_n, \dots$  i.i.d. (independent identical distributed)  $\mu = E[\xi]$ ,  $\sigma^2 = Var[\xi]$ .

**(Strong) Law of Large Numbers** Let  $\xi_1, \dots, \xi_n, \dots$  i.i.d., then  $(\xi_1 + \dots + \xi_n)/n \xrightarrow{(a.s.)-P} \mu$ .

**Borel-Cantelli lemma:** 1.  $\sum_{n=1}^{\infty} P(A_n) < \infty$  then  $P(A_n - i.o.) = 0$

2.  $\sum_{n=1}^{\infty} P(A_n) = \infty$ ,  $A_1 \perp A_2 \perp \dots$ , then  $P(A_n - i.o.) = 1$

( $A_n$  i.o. - infinitely often)  $\{\omega \in \Omega | \#\{n \geq 1 : \omega \in A_n\} = \infty\}$

**Transformations:**  $x = (x_1, \dots, x_n)$   $y = (y_1, \dots, y_n)$ ,  $i = 1, \dots, n$   $y_i = g_i(x_1, \dots, x_n)$ ,

$$h_i(g_i(x)) = x_i, \quad J_g(x) \neq 0, \quad J_g(x) = \partial g_i(x) / \partial x_j \Big|_{i,j=1,\dots,n}, \quad J_h(y) = \partial h_i(y) / \partial y_j \Big|_{i,j=1,\dots,n}$$

$J_g(x)J_h(y)=1$ ,  $X=(X_1, \dots, X_n)$  r.v.,  $Y_i=g_i(X_1, \dots, X_n)$ , then

$$f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = f_{X_1, \dots, X_n}(x_1, \dots, x_n) |J_g(x)|^{-1} = f_{X_1, \dots, X_n}(h_1(y), \dots, h_n(y)) |J_h(y)|$$

**Change of variables in a Lebesgue Integral**  $\int_{\xi^{-1}(A)} g(\xi(\omega)) dP(\omega) = \int_A g(x) dP_\xi(x)$

**Conditional expectation** Let  $E[\xi] < \infty$ ,  $\hat{\xi} = E[\xi | \mathcal{A}]$  if  $\int_A \xi(\omega) dP(\omega) = \int_A \hat{\xi}(\omega) dP(\omega)$  for

all  $A \in \mathcal{A}$ , where  $\hat{\xi}$  is  $\mathcal{A}$ -measurable

**Martingales** Let  $F_1 \subseteq F_2 \subseteq \dots$ ,  $\xi_n$  be a  $F_n$ -measurable.  $(\xi_n, F_n)_{n \geq 1}$  is a (sub)martingale if  $E(\xi_{n+1} | F_n) = (\geq) \xi_n$ .

**Doob's martingale convergence theorem** Let  $(\xi_n, F_n)_{n \geq 1}$  be a submartingale,

$\sup_{n \geq 1} E(\xi_n^+) < \infty$ . Then  $\exists \xi = \lim \xi_n$ ,  $|\xi| < \infty$  a.s.

**Levi's martingale convergence theorem** Let  $E[|\eta|] < \infty$ ,  $\xi_n = E[\eta | F_n]$ ,  $F_\infty = \sigma(\cup_{n \geq 1} F_n)$ .

Then  $\exists \xi_\infty = \lim \xi_n$ ,  $\xi_\infty = E[\eta | F_\infty]$ .

**Reversed Martingales** Let  $E[\xi_{-n} | F_{-m}] = \xi_{-m}$ ,  $F_{-m} \subseteq F_{-n}$  for  $0 \leq n \leq m$ . Then  $\xi_n$  converge a.s. and in  $L^1$  to a limit  $\xi$ .

**Kolmogorov's Trees-Series Theorem** Let  $(\xi_n)$  be a sequence of independent random variables. Then  $\sum \xi_n$  converges almost surely if and only if for some (for every in "only" part)

$K > 0$ , the following three properties hold: 1.  $\sum_n P(|\xi_n| > K) < \infty$  2.  $\sum_n E(\xi_n^K)$

converges 3.  $\sum_n \text{Var}(\xi_n^K) < \infty$  where  $\xi_n^K = \xi_n$  if  $|\xi_n| \leq K$ ,  $\xi_n^K = 0$  if  $|\xi_n| > K$

**Kolmogorov's zero-one law** Let  $A_1^\perp A_2^\perp \dots$ ,  $G = \bigcap_{n \geq 1} \sigma(A_n, A_{n+1}, \dots)$ ,  $A \in G$ . Then  $P(A) = 0$  or 1.

**Brownian motion** **Def.1.**  $G$  is a centered Gaussian process if  $(G(t_1), \dots, G(t_n))$  is a centered Normal random variable in  $R^n$  for all  $0 \leq t_1 < \dots < t_n$ .

$Q(s, t) = E[G(s), G(t)]$  - covariance function of  $G$ .

**Def.2.** The Wiener process (Brownian motion)  $W(t)$  is characterized by three properties

1.  $W(0) = 0$ , 2.  $0 \leq t_1 < \dots < t_n$  ( $W(t)$  has independent increments)

3.  $W_t - W_s \sim \mathcal{N}(0, t-s)$  (for  $0 \leq s < t$ ).

$W(t)$  is a Brownian motion if and only if  $W$  is a centered Gaussian with  $Q(s, t) = \min(s, t)$ .  
 $W(ct)/\sqrt{c} \sim W(t)$  - scaling properties