The aim is to put some basic facts from other modules here that you can quote. They should be sufficient to be able to answer all example sheet and exam questions. I have started and I will add items as things arise that we need during the lectures. I hope many, or even most, of them will appear familiar.

I am including definitions and results but no proofs. I will try and point out in lectures where we are using these facts.

I have assumed that everyone knows about sigma fields (this is now called a sigma algebra in many books), measurable functions, and the basic defining properties of a measure. Evans has a quick summary of these measure theory definitions.

1 Foundations

F1 A random variable X is a measurable function $X : (\Omega, \mathcal{F}, P) \to (E, \mathcal{E})$. Here $(\Omega, \mathcal{F}, P)$ is a probability space (that is P is a non-negative measure with $P(\Omega) = 1$); also here X takes values in a space $(E, \mathcal{E})$. Most often for us $E = \mathbb{R}$ with the Borel sigma field, whereupon X a called a real variable.

F2 The sub-sigma field $\sigma(X) := \{X^{-1}(A) : A \in \mathcal{E}\}$ is called 'information generated by X'.

F3 A stochastic process, for us, is a collection of random variables $(X_t : t \in I)$ all defined on the same probability space. We usually we have $I = [0, \infty)$ or $I = [0, T]$. The sub-sigma field $\sigma(X_t : t \in I)$ is defined as the smallest sigma field containing $\sigma(X_t)$ for all $t \in I$.

F4 The law of an $(E, \mathcal{E})$ valued random variable is the probability defined by $Q(A) = P(X^{-1}(A))$ for $A \in \mathcal{E}$. The *finite dimensional distributions* of a stochastic process $(X_t : t \in I)$ are the laws of the vectors $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ as you vary $t_1, \ldots, t_n \in I$ and $n \geq 1$.

F5 We will meet convergence of random variables $X_n \to X$ in the various senses that probabilists use. The two commonest ways are: $X_n \to X$ almost surely means $P[\{\omega : X_n(\omega) \to X(\omega)\}] = 1$; $X_n \stackrel{L^2}{\to} X$ means $\|X_n - X\|_2 = (E[|X_n - X|^2])^{1/2} \to 0$ as $n \to \infty$. These are not quite equivalent, but they are in the same spirit! Quite different is convergence in distribution: $X_n \stackrel{D}{\to} X$ means that for any bounded continuous F the expectations $E[F(X_n)] \to E[F(X)]$ as $n \to \infty$. This is what you get when you state the central limit theorem. If $X_n \to X$ almost surely or $X_n \stackrel{L^2}{\to} X$ then it follows that $X_n \stackrel{D}{\to} X$. Lots to chat about there but let's see what we need.

2 Independence

Everything here is set in a single probability space $(\Omega, \mathcal{F}, P)$.

I1 Two events $A, B \in \mathcal{F}$ are called *independent* if $P(A \cap B) = P(A)P(B)$.

I2 Two sub sigma fields $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{F}$ are called *independent* if $P(A_1 \cap A_1) = P(A_1)P(A_2)$ for all $A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2$.

I3 Two random variables X and Y are independent means that the sigma fields $\sigma(X)$ and $\sigma(Y)$ are independent. Two stochastic processes $(X_i : i \in I)$ and $(Y_j : j \in J)$ are independent means that the sigma fields $\sigma(X_i : i \in I)$ and $\sigma(Y_j : j \in J)$ are independent.

I4 A sigma field $\mathcal{F}$ is generated by a collection $\mathcal{A}$ if it is the smallest sigma field containing $\mathcal{A}$. A π-system $\mathcal{A}$ is a collection of subsets of Ω that is closed under finite intersection.

Lemma: Suppose two sigma fields $\mathcal{F}_1$ and $\mathcal{F}_2$ are generated by π-systems $\mathcal{A}_1$ and $\mathcal{A}_2$; then you only need to check $P(A_1 \cap A_1) = P(A_1)P(A_2)$ for all $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$ to ensure $\mathcal{F}_1$ and $\mathcal{F}_2$ are independent.

I5 If two sigma fields $\mathcal{F}, \mathcal{G}$ are independent and the two variables X is $\mathcal{F}$ measurable and Y is $\mathcal{G}$ measurable then we expect the usual splitting $E[XY] = E[X]E[Y]$. This is always true if X and Y are non-negative, or if they are both bounded. Another simple sufficient condition is that $E[X^2]$ and $E[Y^2]$ are finite.

The proofs all use the *measure theory machine*: start with X, Y simple variables, that is a sum if indicators $\sum_{i=1}^{N} c_i I(A_i)$. Measure theory explores how to approximate general measurable variables by simple ones. Take increasing limits to show the splitting holds for non-negative case. Take bounded approximations for other cases...

3 Gaussian World

G1 A real Gaussian variable X with a density $e^{-x^2/2}/\sqrt{2\pi}$ on $\mathbb{R}$ and it's law is denoted as $N(0,1)$. The expectation $\mathbb{E}[\exp(\theta X)] = \exp(\theta^2/2)$ for all $\theta \in \mathbb{C}$.

G2 Then the variable $\mu+\sigma X$ is Gaussian and its law is denoted $N(\mu,\sigma^2)$. The collection $(N(\mu,\sigma^2) : \mu \in \mathbb{R}, \sigma^2 \geq 0)$ are the family of Gaussian distributions.

G3 A random vector $(X_1, X_2, \ldots, X_N)$ is called Gaussian if for any real $\lambda_1, \ldots, \lambda_N$ the sum $\lambda_1 X_1 + \ldots + \lambda_N X_N$ has a Gaussian distribution.

A stochastic process $(X_t : t \in I)$ is called Gaussian if all vectors $(X_{t_1}, \ldots, X_{t_n})$ for any n and (t_i) are Gaussian.

G4 Lemma: if $X_1 \stackrel{D}{=} N(\mu_1, \sigma_1^2)$ and $X_2 \stackrel{D}{=} N(\mu_2, \sigma_2^2)$ are independent then
$X_1 + X_2 \stackrel{D}{=} N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

G5 The finite dimensional distributions of a Gaussian stochastic process $(X_t : t \in I)$ are determined by the means $E[X_t]$ and correlations $E[X_s X_t]$ for $s, t \in I$. This an exercise (use transforms) that is well known to all statisticians.

G6 Lemma: Take a pair of jointly Gaussian variables (X, Y). Then X and Y are independent if and only if $E[XY] = E[X]E[Y]$.

The analogues of this is true for Gaussian vectors $X = (X_1, \ldots, X_{N+M})$: then $(X_1, \ldots, X_N)$ is independent of $(X_{N+1}, \ldots, X_{N+M})$ provided $E[X_i X_j] = E[X_i]E[X_j]$ for all $i \in \{1, \ldots, N\}$ and $j \in \{N+1, \ldots, N+M\}$.

G7 'The limit of Gaussians must be Gaussian'. For example if $X_n \stackrel{D}{=} N(\mu_n, \sigma_n^2)$ then $X_n \stackrel{D}{\to} X$ if and only if $\mu_n \to \mu$, $\sigma_n^2 \to \sigma^2$ and $X \stackrel{D}{=} N(\mu, \sigma^2)$. Not tricky to show using transforms.

4 Transforms

T1 For random variables $X = (X_1, \ldots, X_n)$ with values in $\mathbb{R}^n$, probabilists often use the Fourier transform $\phi_X(\theta) = E[\exp(i(\theta_1 X_1 + \ldots + \theta_n X_n))]$, for $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$ which they call the characteristic function. It is characteristic, meaning the values $(\phi(\theta) : \theta \in \mathbb{R}^n)$ determine the law of X.

T2 Transforms mesh well with independence: two variables $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ and $Y = (Y_1, \ldots, Y_m) \in \mathbb{R}^m$ are independent if and only if

$$\phi_{X,Y}(\theta, \varphi) = E[\exp(i(\theta_1 X_1 + \ldots + \theta_n X_n)) \exp(i(\varphi_1 Y_1 + \ldots + \varphi_m X_m))] = \phi_X(\theta)\phi_Y(\varphi)$$

for all $\theta \in \mathbb{R}^n$, $\phi \in \mathbb{R}^m$.

T3 Transforms mesh well with convergence in distribution. Real variables X_n converge in distribution to a limit X means that $E[F(X_n)] \to E[F(X)]$ as $n \to \infty$ for all bounded continuous $F : R \to \mathbb{R}$. It is enough to check this for suitable sets of functions, and transforms are enough - thus if $\phi_{X_n}(\theta) \to \phi_X(\theta)$ for all θ this implies convergence in distribution. The analogous statement for $\mathbb{R}^d$ valued variables, that is random vectors, is true.

5 Brownian Motion

We can define Brownian motion as a real stochastic process $(W_t : t \in [0,1])$ satisfying

(i) $W_0 = 0$ and $t \to W_t$ is continuous, almost surely.

(ii) $W_t - W_s$ is independent of $(W_r : r \leq s)$ for all $0 \leq s \leq t$.

(iii) $W_t - W_s \overset{D}{=} N(0, t-s)$ for all $0 \leq s \leq t$.

An equivalent definition is to replace (ii) and (iii) by

(ii)' $(W_t : t \geq 0)$ is a Gaussian process.

(iii)' $E[W_t] = 0$ and $E[W_s W_t] = \min\{s,t\}$ for all $0 \leq s \leq t$.

BM1 Brownian motion exists.

BM2 New from old: the following transformations of a BM are also a BM

i. **Reflection:** $X_t : -W_t$ for $t \geq 0$.
ii. **Time shift:** $X_t := W_{t_0+t} - W_t$ for a fixed t_0 and $t \geq 0$.
iii. **Scaling:** $X_t : cW_{t/c^2}$ for a fixed $c > 0$ and $t \geq 0$.
iv. **Time inversion:** $X_t : tW_1/t$ for $t > 0$ and $X_0 = 0$.

BM3 Recurrence: $\sup_{t\geq 0} W_t = +\infty$ and $\inf_{t\geq 0} W_t = -\infty$ almost surely.

BM4 Reflection formula: $P[\sup_{s\leq t} W_s \geq a] = 2P[W_t \geq a]$.

BM5 Growth of sample paths: $\lim_{t\to\infty} |W_t|/t^p = 0$ for $p > 1/2$ and $\limsup_{t\to\infty} |W_t|/t^p = \infty$ for $p \in [0, \frac{1}{2}]$ almost surely.