- **Continuous pdf:**
  $f_X(x) : \mathcal{X} \to [0, \infty)$, $f_X(x) = \lim_{\epsilon\to 0} \frac{1}{\epsilon} \cdot \Pr[x \le X \le x+\epsilon]$
- **Continuous cdf:**
  $F_x(x) := \Pr[X \le x]$ where $\Pr[a \le X \le b] = \int_a^b f_x(x)\,\mathrm{d}x$
- **Uniform distribution on $\mathcal{X} = [u, v]$:**
  $f_X(x) = 1/(v-u)$ for $x \in \mathcal{X}$
- **Normal distribution:**
  $f_X(x) = (2\pi)^{-1/2} \exp(-x^2/2)$
- **Dirac delta property:**
  $\int_{-\infty}^{\infty} \mathrm{d}x\, b(x)\delta(x-a) = b(a)$
- **Expectation value of $g(X)$ where $X \in \mathcal{X}$:**
  $\mathbb{E}[g(X)] = \sum_{x\in\mathcal{X}} \Pr[X = x] g(x)$, k'th moment of $X = \mathbb{E}[X^k]$
- **Statistical distance of $X, Y \in \mathcal{X}$:**
  $\Delta(X, Y) = \frac{1}{2}\sum_{x\in\mathcal{X}} |\mathbb{P}(x) - \mathbb{Q}(x)|$
- **Covariance matrix K:**
  $K_{i,j} = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \cdot \mathbb{E}[X_j]$
  Zero covariance ($K_{1,2} = K_{2,1} = 0$) does not imply X_1 and X_2 are independent.
- **Marginal distribution for X when $(X, Y) \sim \mathbb{P}$:**
  $\Pr[X = x] = \sum_y \mathbb{P}(x, y)$
- **Conditional probability for $(X, Y) \sim \mathbb{P}$:**
  $\Pr[X = x|Y = y] = \frac{\Pr[X=x, Y=y]}{\Pr[Y=y]} = \frac{\mathbb{P}(x,y)}{\mathbb{P}_2(y)}$
- **(Shannon) Entropy rules:**
  1. **Additivity**: inf of a set of indep. RVs must be the sum of indiv. inf. contents
  2. **Sub-additivity**: Total inf. content of two jointly distrib. RVs cannot exceed sum of seperate infs.
  3. **Expansibility**: Adding extra outcome of prob. 0 does not affect inf.
  4. **Normalization**: The distrib $(1/2, 1/2)$ has inf. of 1 bit.
  5. The distrib $(p, 1-p)$ for $p \to 0$ has zero inf.
- **Shannon entropy:**
  Lower bound on the avg length of the shortest desc of X.
  $H(X) = \sum_{x\in\mathcal{X}} p_x \log_2 \frac{1}{p_x}$
  $H(X,Y) = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p_{xy} \log_2 \frac{1}{p_{xy}}$
- **Renyi entropy:**
  $H_\alpha(X) = \frac{-1}{\alpha-1} \log \sum_{x\in\mathcal{X}} p_x^\alpha$
- **Binary entropy function: 2 outcomes with prob. p and $1-p$:**
  $h(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$
- **Differential entropy for continuous RV $X \sim \rho$:**
  $h_{\text{diff}}(X) = -\int \mathrm{d}x\, \rho(x) \log \rho(x) = \mathbb{E}_x \log \frac{1}{\rho(x)}$
- **Relative entropy (Kullback-Leibler distance):**
  $D(\mathbb{P}||\mathbb{Q}) = \sum_{x\in\mathcal{X}} \mathbb{P}(x) \log \frac{\mathbb{P}(x)}{\mathbb{Q}(x)}$
- **Entropy of jointly distrib. RVs: $H(X,Y)$ or $H(XY)$:**
  $H(X,Y) = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p_{xy} \log \frac{1}{p_{xy}}$
- **Conditional entropy:**
  $H(X|Y) = \mathbb{E}_y[H(X|Y = y)] = -\sum_{x\in\mathcal{X}} p_x \sum_{y\in\mathcal{Y}} p_{x|y} \log p_{x|y}$
  $H(X|Y) = H(X,Y) - H(Y)$
- **Mutual information:**
  $\mathbf{I}(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$
  $\mathbf{I}(X;Y) = H(X,Y) - H(X|Y) - H(Y|X)$
  $\mathbf{I}(X;Y) = H(X) + H(Y) - H(X,Y)$
  $\mathbf{I}(X;Y|Z) = \mathbb{E}_z \mathbf{I}(X|Z = z; Y|Z = z)$
- **Min entropy:**
  $H_{\min}(X) = -\log \max_{x\in\mathcal{X}} p_x = -\log p_{\max}$
  $H_{\min}(X|Y) = -\log \mathbb{E}_y \max_{x\in\mathcal{X}} p_{x|y}$
- **Linear binary codes:**
  Maps k-bit msg x to n-bit ($n > k$) codeword $c_x \in \mathcal{C}$. Perceived string $z = c_z \oplus e$. Minimum distance of code: $d = \min_{c,c'\in\mathcal{C}} \text{HammingWeight}(c \oplus c')$. Receiver determines which $c_{\hat{x}}$ is closest to z and decodes it into $\hat{x}$. Error correcting capability $t = \lfloor \frac{d-1}{2} \rfloor$.
- **Generator (G is $k \times n$) and parity check (H is $(n-k) \times n$) matrix:**
  $c_x = xG$. $G = (\mathbf{1}_k|A)$. $H = (-A^T|\mathbf{1}_{n-k})$. $GH^T = 0$, $cH^t = 0$.
  All k rows of G are linearly independent.
- **Syndrome decoding ($s(z) \in \{0,1\}^{n-k}$):**
  $s(z) = zH^T = (c_x + e)H^T = eH^T$
  Syndrome depends only on the error pattern, not on the message.
- **Hamming bound: Binary code of length n that can correct t errors:**
  $2^k \le 2^n / \sum_{j=0}^{t} \binom{n}{j}$. Approx $\log n$ bits of redundancy per bit error.
- **Channel capacity:**
  Inf. content error free: $k \le \mathbf{I}(C; Z)$
  BSC capacity (per bit): $\frac{k}{n} \le \mathbf{I}(C_j; Z_j) = H(Z_j) - H(Z_j|C_j)$
  This is called the BSC code rate: BSC CODE RATE $\le 1 - h(\epsilon)$
  Following the rule of thumb: $h(\epsilon) = -\epsilon \log \epsilon + \mathcal{O}(\epsilon)$
- **Uniformly random bits from continuous source:**
  TODO
- **Randomness sources:**
  Ring oscilators (odd number of inverters causes jitter in period, gaussian), noisy resistors (no voltage applied, noise amplitude gaussian distribution), radioactive decay (poisson)
- **The von Neumann alg:**
  Given (b_1, b_2), if $b_1 = b_2$ then no output, else output b_1.
- **Piling-up lemma:**
  Let $X_1, ..., X_n \in \{0,1\}$ be independent with biases $\Pr[X_i = 1] - \Pr[X_i = 0] = \alpha_i$. Construct $Y = X_1 \oplus X_2 \oplus ... \oplus X_n$. The bias of Y is $\Pr[Y = 1] - \Pr[Y = 0] = (-1)^{n-1} \prod_{i=1}^{n} \alpha_i$. Thus by xoring many bits together the bias gets reduced.
- **Resilient function:**
  A function $\Psi : \{0,1\}^n \to \{0,1\}^m$ is (n, m, t)-resilient of, for any t coords $i_1, ..., i_t \in [n]$, any $a_1, ..., a_t \in \{0,1\}$ and any $y \in \{0,1\}^m$ it holds that: $\Pr[\Psi(X) = y|x_{i_1} = a_1, ..., x_{i_t} = a_t] = 2^{-m}$
  i.e.: Knowledge of t values of the input does not give inf. that would help in guessing the output. ECC example ($[n, k, d]$ code): $\Psi : \{0,1\}^n \to \{0,1\}^k$. $\Psi = xG^T$. Then Ψ is an $(n, k, d-1)$-resilient fun.
- **Strong extractor Ext : $\{0,1\}^n \times \{0,1\}^* \to \{0,1\}^l$:**
  Takes n-bit string X and randomness R and outputs an l-bit string ($l < n$). $Z = \text{Ext}(X, R)$. Ext is a strong extractor for source min-entropy m, output length l and nonuniformity ϵ if for all distrib of X with $H_\infty(X) \ge m$ it holds that $\Delta(ZR; U_l R) \le \epsilon$. U_l is an RV uniform on $\{0,1\}^l$. Also true: $\mathbb{E}_r \Delta(Z|R = r; U_l) \le \epsilon$
- **Extractable randomness:**
  TODO
- **Universal hash functions:**
  Let $\mathcal{R}$, $\mathcal{X}$ and $\mathcal{T}$ be finite sets. Let $\{\Phi_r\}_{r\in\mathcal{R}}$ be a family of hash functions from $\mathcal{X}$ to $\mathcal{T}$. The family is called universal iff, for R drawn uniformly from $\mathcal{R}$, it holds that: $\Pr[\Phi_R(x) = \Phi_R(x')] \le 1/|\mathcal{T}|$. It's called η-almost universal if it holds that: $\Pr[\Phi_R(x) = \Phi_R(x')] \le \eta$ (for all $x, x' \in \mathcal{X}$ with $x \ne x'$).
- **Leftover hash lemma:**
  TODO
- **Noisy broadcast channel and no return channel:**
  Secret capacity C_s of the broadcast channel $P_{Y\,Z|X}$ can be bounded: $C_s(P_{Y\,Z|X}) \ge \max_{P_X}[\mathbf{I}(X;Y) - \mathbf{I}(X;Z)] = \max_{P_X}[H(X|Z) - H(X|Y)]$. Condition: if Eve's reception quality is better than Bob's ($\mathbf{I}(X;Y) < \mathbf{I}(X;Z)$) then the secrecy capacity is zero. Secrecy capacity of BSC with error rates ϵ and δ is: $h(\delta) - h(\epsilon)$ if $\delta > \epsilon$, and 0 otherwise.

- **Noisy broadcast channel plus public return channel:**
$\hat{C}_s(P_{Y\,Z|X}) \leq \min\{\max_{P_X} \mathbf{I}(X;Y), \max_{P_X} \mathbf{I}(X;Y|Z)\}$
$\hat{C}_s(\epsilon,\delta) = h(\epsilon * \delta) - h(\epsilon)$.
- **Satellite scenario:**
TODO
- **PUF types and their properties:**
General properties:
  - The object can be subjected to a large number of diff challenges that yield an unpredictable response
  - The object is very hard to clone physically
  - Mathematical modeling of the challenge-response physics is very difficult
  - Opaqueness: It is hard to characterize the physical structure of the object in a non-destructive way.

  Types:
  - **Coating PUF**: random layor of conductors and insulators: probes result in binary string. Used for secure key storage.
  - **Optical PUF**: 3D optical structure produces speckle pattern. Challenge: props of laser beam: angle of incidence, focal dist.
  - **Silicon PUF**: variations in IC from manufacturing. Challenge: pulsed time signal to certain part. Response: delay times of various wires and logic devices.
  - **SRAM PUF**: Undefined state of RAM cells. Challenge: memory address, response: returned start-up values.
  - **Randomly positioned glass fibers**: Challenge: ordinary beam of light lighting up part of the layer. Response: Certain fibers light up.

  Uncontrolled PUF: reader interacts directly with PUF structure, trusted reader. Controlled PUF (CPUF): interaction through a *control layer*, PUF and *cl* are bound together, seperation will damage the PUF. Result: attacker has no direct access to PUF. Example: secure key storage where control layer performs zk-protocol to prove knowledge of the key (called *Physically Obscured Key (POK)*).
- **PUF math:**
Information revealed by a noisy meassurement outcome U' where $m \in \mathcal{M}(m(K) = U)$: $\mathbf{I}(U';K) = \mathbf{I}(U';U)$. Noiseless case: $\mathbf{I}_m$

  Meassurable entropy of PUF (space $\mathcal{K}$ and $\mathcal{M}$):
$\mathbf{I}^{\text{meas}}_{\mathbb{P}\mathcal{M}} = \max_{m\in\mathcal{M}} H(m(k))$

  Security param of bare PUF: min num of C-R meassurements required to reveal all measurable info of the PUF: $S_{\mathbb{P}\mathcal{M}_0}$
- **Fuzzy extractor:**
Gen and Rep algorithms. $(S_x, W_x) = Gen(X)$. $S'_x = Rep(X', W_x)$ Must satisfy the following properties:
  - **Correctness**: The prob that $S'_x = S_x$ must be close to 1.
  - **Security**: The RV S_x must be close to uniform, given knowledge of W_x.
- **Secure Sketch** (for discrete src space $\mathcal{X}$):
$SS : x \mapsto w_x$, $Rec : (x', w_x) \mapsto \hat{x}$ with:
  - **Correctness**: The prob that $\hat{X} = X$ must be close to 1.
  - **Security**: X given W_X must have high entropy.
- **When to use FE and SS:**
FE: reliably extract a cryptographic key from noisy data. SS: reliably extract a string with sufficient (min-)entropy. Easier to construct SS than FE, in general: SS extracts more (min-)entropy than a FE from the same source.
- **Code offset method (COM):**
Enroll (Gen): $s \in \{0,1\}^k$, $c_s = Enc(s)$. $w = c_S \oplus x$. Output s as secret and w as helper data. Reconstruct (Rep): $\hat{s} = Dec(x' \oplus w)$.
- **Zero leakage FE (for continuous RV) based on partitions**
$\Pr[S = s|W = w] = 1/n$. Enroll: determine which partition the meassured val x is located in: $\mathcal{A}_{ij}$. Set $s_x = i$, $w_x = j$. Reconstruct: X' is meassured, determine for which s' the interval $\mathcal{A}_{s',w_x}$ is closest to x'. This s' is the reconstructed key. $H_\infty(S|W = w) = H_\infty(S|W) = \log n$.
- **Helper data schemes for specific PUF types:**
TODO
- **Distance bounding principles (and fraud types):**
Mafia fraud: challenge is relayed to different location where a legit. device is tricked into giving a response, response is relayed back to verifier. **Terrorist fraud**: legit. device cooperates with attacker, does not have to follow protocol, can share everything except long-term auth secrets.
(Light travels about 300m every μs). If a device repeatedly correctly respons to an *unpredictable* challenge within time Δt, then the location where the response is computed cannot be further away than $x = c\Delta t/2$. Max time for resp. to arrive: $t_{\text{max}} = 2\frac{x_{\text{max}}}{c} + t_{\text{slack}}$.
- **Brands-chaum protocol:**
commit, rapid bit phase, sign phase. No link between phases, mafia not possible (timing meassured), terrorist is possible.
- **Swiss knife protocol**
Rapid phase uses $R0_i = Z0_j$ and $R1_i = Z1_j$ to determine response r_i. $Z1 = Z0 \oplus x$ so attacker cannot perform rapid phase without knowledge of the key. Therefore terrorist and mafia are not possible. **Analog impl.** Similar to swiss knife but with LP and HP filters.
- **Linear algebra:**
Complex conjugate of $a + bi$ is $a - bi$. conjugate of $\rho e^{i\phi}$ is $\rho e^{-i\phi}$. Hermitian conjugate of a complex number is its complex conjugate: $a^\dagger = a^*$. On a matrix: Hermitian conjugate is transpose followed by complex conjugate. The Hermitian conjugate of $|\psi\rangle$ is $\langle\psi|$.
- **Quantum stuff:**
Meassurement destroys state information.
Time evolution of a quantum system can be represented as a unitary operator acting on a starting space. A unitary op U is defined as: $UU^\dagger = \mathbf{1}$ and $U^\dagger U = \mathbf{1}$. Norm of a vec is preserved: For $|\psi'\rangle = U|\psi\rangle$, then the norm is $\langle\psi'|\psi'\rangle = \langle\psi|U^\dagger U|\psi\rangle = \langle\psi|\psi\rangle$.

  Tensor product: $\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \otimes \begin{pmatrix}\gamma\\ \delta\end{pmatrix} = \begin{pmatrix}\alpha\gamma\\ \alpha\beta\\ \beta\gamma\\ \beta\delta\end{pmatrix}$.

  For qubits, tensor is omitted: $|0\rangle \otimes |1\rangle = |01\rangle$
- **No cloning theorem:**
Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. let $|\psi\rangle \in \mathcal{H}_1$ and $|e\rangle \in \mathcal{H}_2$, where e is known and ψ is unknown. Then there does not exist a unitary operator U_e acting on $\mathcal{H}_1 \otimes \mathcal{H}_2$ satisfying $U_e|\psi\rangle \otimes |e\rangle = |\psi\rangle \otimes |\psi\rangle$ for all ψ
- **Quantum readout of PUFs:**
One round of the protocol:
  1. The verifyer choses a random challenge ψ. He prepares a particle in state $|\psi\rangle$ and sends the particle to the prover.
  2. The prover lets the particle interact with the PUF. This results in a state $|\omega\rangle = R|\psi\rangle$. He sends the particle back to the verifyer.
  3. The verifier does a meassurement $|\omega\rangle\langle\omega|$ on the particle. If the outcome is 1 then the prover has passed this round.