- 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:
$H(X) = \sum_{x\in\mathcal{X}} p_x \log_2 \frac{1}{p_x}$
- 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
- 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 $\text{Ext} : \{0,1\}^n \times \{0,1\}^* \to \{0,1\}^l$:
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 $\Delta(ZR; U_l R) \le \epsilon$.
- Universal hash functions:
- Leftover hash lemma:
- Binary symmetric channel:
- Secret capacity:
- PUF types and their properties:
- PUF applications + attack models:
- PUF entropy and inf. stuff:
- Fuzzy extractor definition:
- When to use FE and SS:
- Zero leakage scheme stuff:
- Helper data scheme:
- Distance bounding principles (and fraud types):
- Brands-chaum protocol:

- Swiss knife protocol
- Analog impl.:
- Quantum stuff: