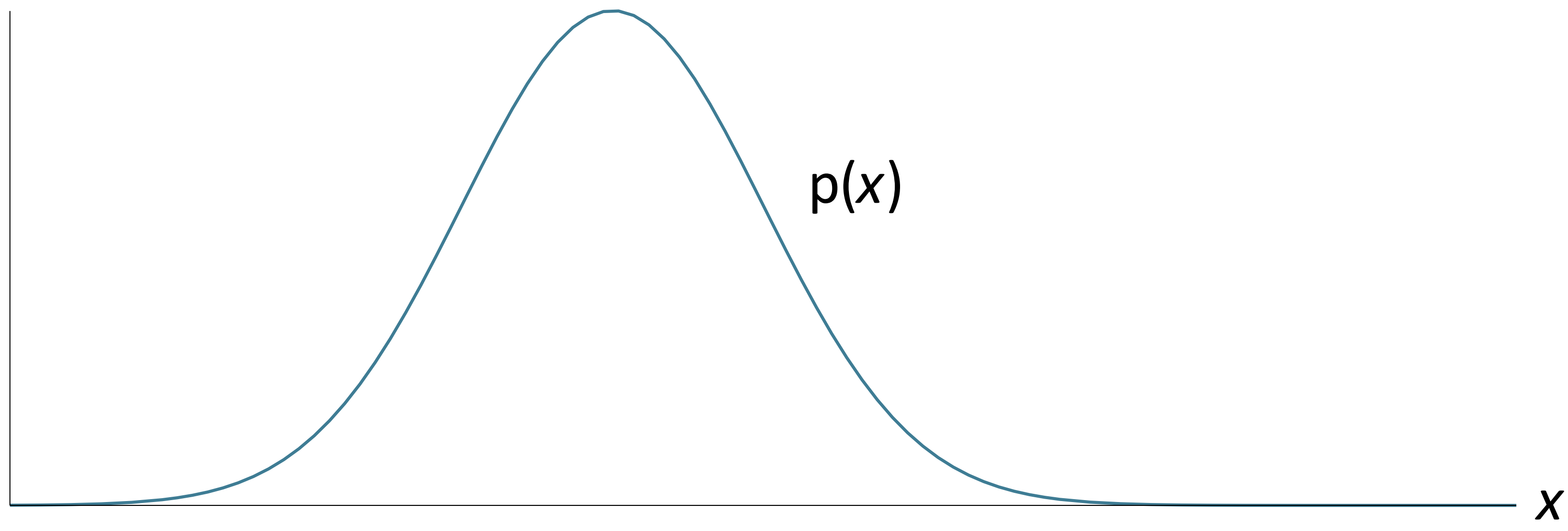
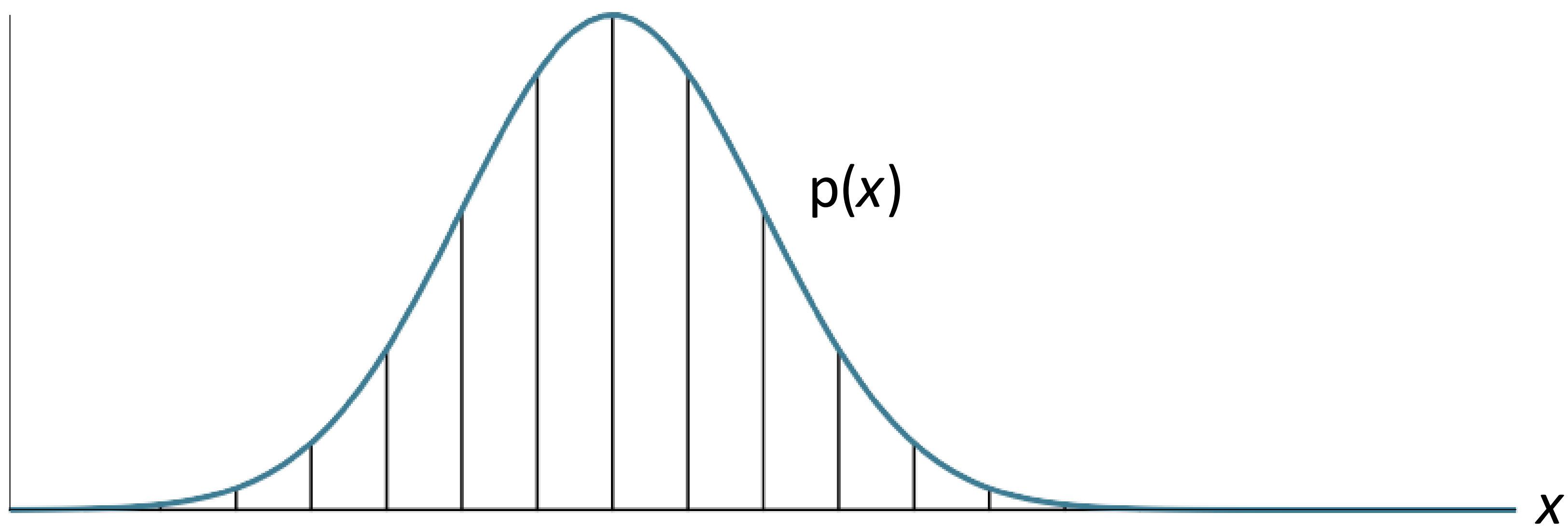
ELEC 515
Information Theory

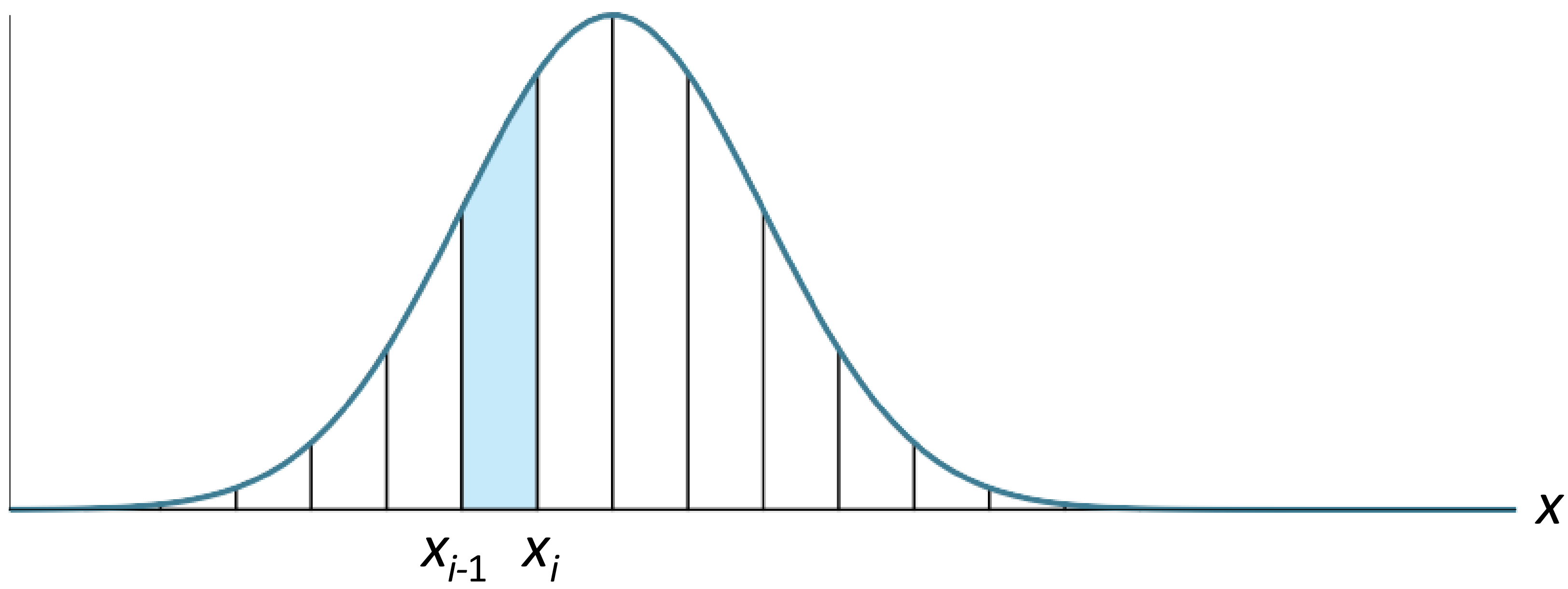
Differential Entropy

- Consider a continuous RV X with probability density function (pdf) p(x) and support S (values for which p(x) > 0)
- We can use X to define a discrete RV

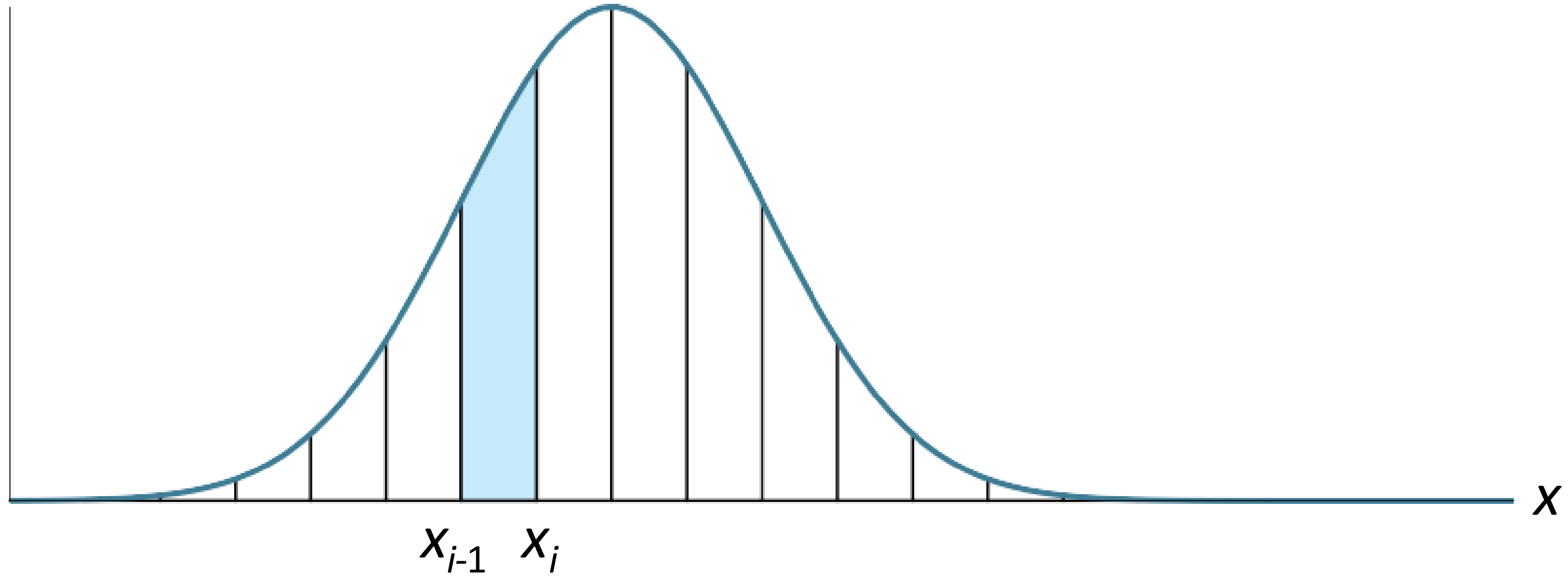

p(x)
x


p(x)
x

Let y_i = $[x_{i\text{-}1}, x_i)$ be a subinterval of width $\Delta = x_i - x_{i\text{-}1}$



Assign to y_i the probability $q_i = \int_{x_{i-1}}^{x_i} \mathrm{p}(x)dx$



- The RV X^Δ whose outcomes are the y_i has entropy

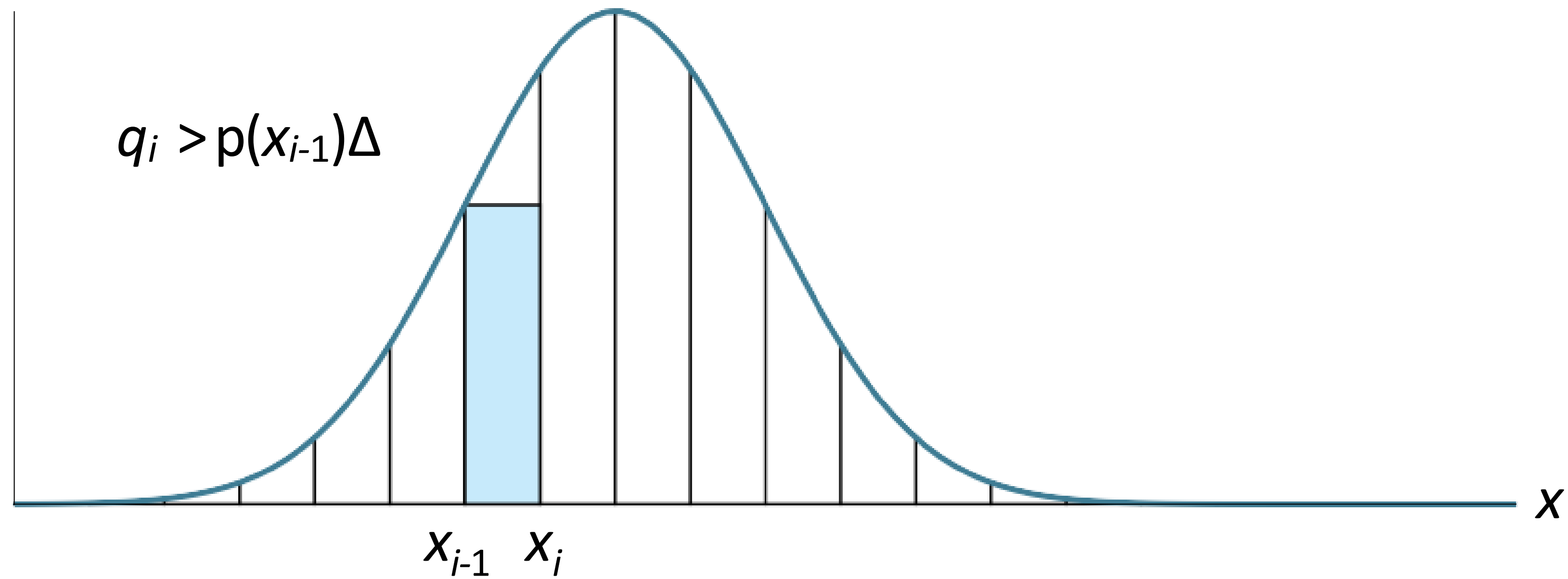
$$H(X^\Delta) = -\sum_i q_i \log q_i$$

- $q_i = \int_{x_{i-1}}^{x_i} \mathrm{p}(x)dx \approx \mathrm{p}(\overline{x_i})\Delta$

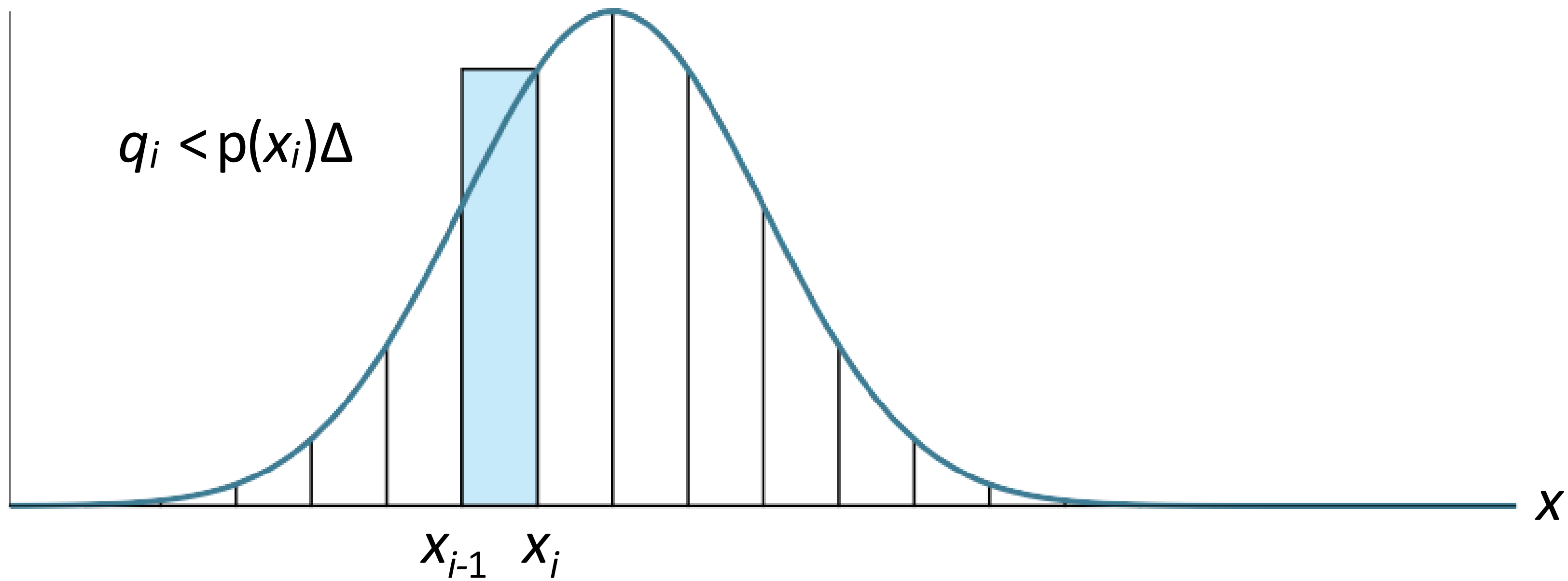
where $\overline{x_i}$ is a point in the subinterval

$y_i = [x_{i-1}, x_i)$

and the approximation gets better as Δ gets smaller

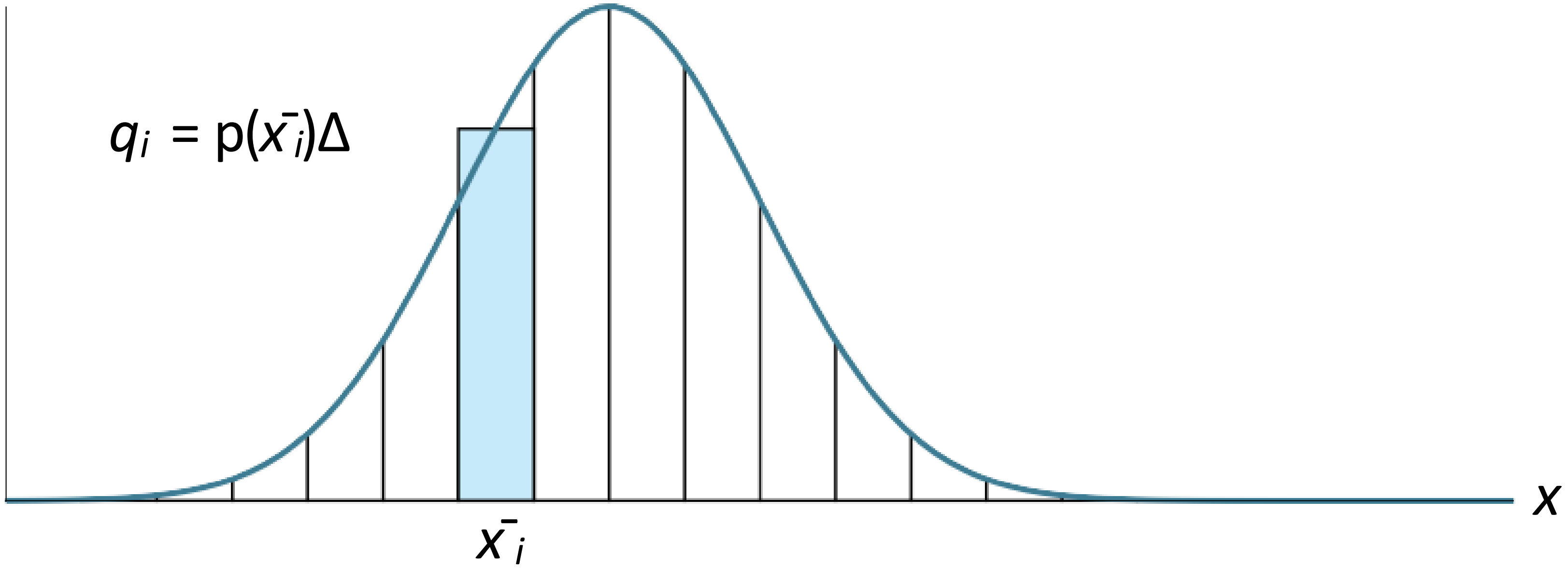


$q_i > p(x_{i-1})\Delta$
x_{i-1}
x_i
x


$q_i < p(x_i)\Delta$
x_{i-1}
x_i
x

From the Mean Value Theorem, if p(x) is continuous we can always pick a value of $\overline{x_i}$ such that

$$\mathrm{p}(\overline{x_i})\Delta = \int_{x_{i-1}}^{x_i} \mathrm{p}(x)dx$$



$$\mathrm{H}(\mathrm{X}^{\Delta}) = -\sum_i q_i \log q_i - \sum_i \mathrm{p}(\overline{x_i})\Delta \log \mathrm{p}(\overline{x_i})\Delta$$

Expanding the log and using

$$\sum_i \mathrm{p}(\overline{x_i})\Delta = \int \mathrm{p}(x)dx = 1$$

gives

$$\mathrm{H}(\mathrm{X}^{\Delta}) = -\sum_i \mathrm{p}(\overline{x_i})\Delta \log \mathrm{p}(\overline{x_i}) - \log\Delta$$

Using the Riemann approximation

$$\sum_i \mathrm{p}(\overline{x_i})\Delta \rightarrow \int \mathrm{p}(x)dx$$

as $\Delta \rightarrow 0$

gives

$$\mathrm{H}(\mathrm{X}^{\Delta}) = -\int \mathrm{p}(x)\log\mathrm{p}(x)dx - \log\Delta$$

as $\Delta \rightarrow 0$

Differential Entropy

The differential entropy of X is defined as

$$\mathrm{h(X)} \triangleq -\int_S \mathrm{p}(x)\log\mathrm{p}(x)dx = \mathrm{E}[-\log\mathrm{p}(x)]$$

where S is the support of X

Then $\mathrm{H(X^{\Delta})} = \mathrm{h(X)} - \log\Delta$

as $\Delta \to 0$

Uniform Distribution

- Consider a random variable distributed uniformly from 0 to *a* so that its density is 1/*a* from 0 to *a* and 0 elsewhere
- Then its differential entropy is

$$\mathrm{h(X)} = -\int_0^a \frac{1}{a}\log\frac{1}{a}\,dx = \log a$$

- Note: For $a < 1$, $\log a < 0$, and the differential entropy is negative. Hence, unlike discrete entropy, differential entropy can be negative

Gaussian Distribution

- pdf $\mathrm{f}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x^2}{2\sigma^2}}$

$$\mathrm{h(X)} = -\int \mathrm{f}(x) \ln \mathrm{f}(x) dx = -\int \mathrm{f}(x) \left[-\frac{-x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right] dx$$

$$= \frac{\mathrm{E[X^2]}}{2\sigma^2} + \frac{1}{2} \ln 2\,\pi\sigma^2 = \frac{1}{2} + \frac{1}{2} \ln 2\,\pi\sigma^2 = \frac{1}{2} \ln e + \frac{1}{2} \ln 2\,\pi\sigma^2$$

$= \frac{1}{2} \ln 2\,\pi e\sigma^2$ nats

- Changing the base of the logarithm gives

$\mathrm{h(X)} = \frac{1}{2} \log_2 2\pi e\sigma^2$ bits

Joint Differential Entropy

- Consider two RVs X and Y with joint pdf p(x,y)
- The joint differential entropy is

$$\mathrm{h(XY)} = -\int \mathrm{p}(x, y) \log \mathrm{p}\,(x, y) dxdy$$

Mutual Information

- The mutual information I(X;Y) between two random variables with joint density f(*x*,*y*) is defined as

$$\mathrm{I}(\mathrm{X};\mathrm{Y}) = \int \mathrm{f}(x,y)\log\frac{\mathrm{f}(x,y)}{\mathrm{f}(x)\mathrm{f}(y)}\,dxdy$$

- I(X;Y) ≥ 0 with equality iff X and Y are independent
- From the definition we have that

$$\begin{aligned} \mathrm{I}(\mathrm{X};\mathrm{Y}) &= \mathrm{h}(\mathrm{X}) - \mathrm{h}(\mathrm{X}|\mathrm{Y}) \\ &= \mathrm{h}(\mathrm{Y}) - \mathrm{h}(\mathrm{Y}|\mathrm{X}) \\ &= \mathrm{h}(\mathrm{X}) + \mathrm{h}(\mathrm{Y}) - \mathrm{h}(\mathrm{XY}) \end{aligned}$$

Relative Entropy

- The relative entropy D(p(X)||q(X)) between two probability densities p(X) and q(X) is defined as

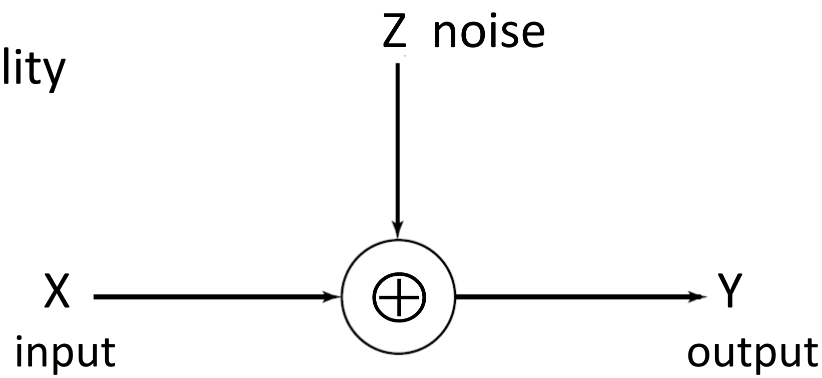
$$\mathrm{D}(\mathrm{p}(\mathrm{X})||\mathrm{q}(\mathrm{X})) = \int \mathrm{p}(x)\log\frac{\mathrm{p}(x)}{\mathrm{q}(x)}\,dx$$

- D(p(X)||q(X)) ≥ 0 with equality iff p(X) = q(X)

BSC Channel Capacity

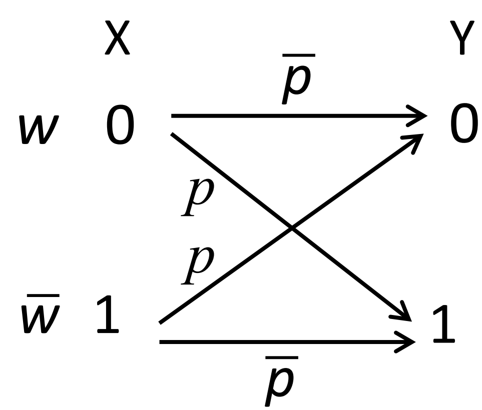
crossover probability

p = Pr(z = 1)



$$\Pr(x=0)=w$$

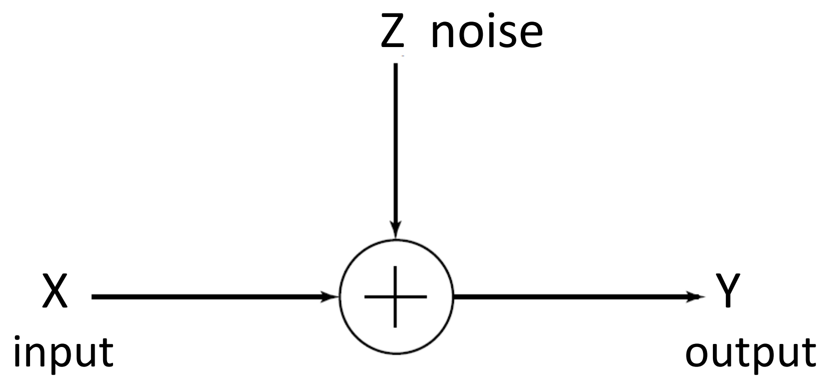
$$\Pr(x=1)=1-w=\bar{w}$$



$$w=\bar{w}=\frac{1}{2}$$

$$C=1-h(p)$$

AWGN Channel Capacity



$$f(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2}\right]$$

AWGN Channel Capacity

$$C = W\log_2\left(1+\frac{P}{N_0 W}\right)$$

$$E = PT \rightarrow P = E_b R_b$$

$$C = W\log_2\left(1+\frac{E_b R_b}{N_0 W}\right)$$

Let $R_b = C$

$$\frac{C}{W} = \log_2\left(1+\frac{E_b}{N_0}\frac{C}{W}\right)$$

$$\frac{E_b}{N_0} = \frac{2^{C/W}-1}{C/W}$$

Bandwidth Efficiency versus SNR