

## Chapter 2:

(Markov's Inequality)

$$P[X \geq t] \leq \frac{E[X]}{t}$$

(Chebyshov's Inequality)

$$P[|X-\mu| \geq t] \leq \frac{\text{Var}(X)}{t} \quad \text{for } t > 0$$

2.1

$$\begin{aligned} a) P[X \geq t] &= \sum_{X=t+1}^n P[X_i \geq t] = P[t+1 \geq t] + P[t+2 \geq t] + P[t+3 \geq t] + \dots \\ &= \frac{E[X]}{t} \end{aligned}$$

$$b) P[|Y-\mu| \geq t] = P[(Y-\mu)^2 \geq c^2] = \sum_{(Y-\mu)^2=c^2+1}^n P[(Y-\mu)^2 \geq c^2] = \frac{E[(Y-\mu)^2]}{c^2} = \frac{\sigma^2}{c^2}$$

2.2

$$a) \Phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}; z \sim N(0, 1)$$

$$\begin{aligned} \phi'(z) &= -\frac{z}{\sqrt{2\pi}} e^{-z^2/2} \quad \text{and} \quad \phi' + z\phi(z) = -\frac{z}{\sqrt{2\pi}} e^{-z^2/2} + \frac{z}{\sqrt{2\pi}} e^{-z^2/2} \\ &= 0 \end{aligned}$$

$$b) \Phi(z) \left( \frac{1}{z} - \frac{1}{z^3} \right) \leq P[Z \geq z] \leq \Phi(z) \left( \frac{1}{z} - \frac{1}{z^3} + \frac{3}{z^5} \right)$$

$$\text{From part a, } \phi'(z) + z\phi(z) = 0 \text{; } z = -\frac{\phi'(z)}{\phi(z)}$$

$$-\Phi(z) \left( \frac{\phi(z)}{\phi'(z)} - \frac{\phi''(z)}{\phi'(z)^3} \right) \leq -\frac{E[z]\phi(z)}{\phi'(z)} \leq -\Phi(z) \left( \frac{\phi(z)}{\phi'(z)} - \frac{\phi''(z)}{\phi'(z)^3} + \frac{3\phi''(z)}{\phi'(z)^5} \right)$$

$$0 \leq \frac{E[z]}{\phi'(z)} + \frac{\phi(z)}{\phi'(z)} - \frac{\phi''(z)}{\phi'(z)^3} \leq -\frac{3\phi''(z)}{\phi'(z)^5}$$

(Markov's Upper Bound)

$$P[(X-\mu) \geq t] = P[e^{\lambda(X-\mu)} \geq e^{\lambda t}] \leq \frac{E[e^{\lambda(X-\mu)}]}{e^{\lambda t}}$$

(Chernoff's Upper Bound)

$$\log P[(X-\mu) \geq t] \leq \inf_{\lambda \in [0, b]} \left\{ \log E[e^{\lambda(X-\mu)}] - \lambda t \right\}$$

2.3.

$$2.3. P[|X|^k \geq \delta^k] = \inf \frac{E[|X|^k]}{\delta^k} = \inf \frac{E[e^{\lambda X}]}{e^{\lambda \delta}}$$

where  $\lambda > 0$

2.4.  $X \sim \mu = E[X]; b > a \Rightarrow X \in [a, b]$

a)  $\psi(\lambda) = \log E[e^{\lambda X}]$

$$\psi(0) = \log(1) = 0$$

$$\psi'(0) = \frac{X e^{\lambda X}}{E[e^{\lambda X}]} = \frac{X}{n} = \mu$$

b)  $\psi''(\lambda) = \frac{X^2 e^{\lambda X} E[e^{\lambda X}] - X e^{\lambda X} E[F(X)e^{\lambda X}]}{E[e^{\lambda X}]^2}$

$$= \frac{X^2}{n^2} - (e^{\lambda X})' \cdot E[F(X)]$$

$$= E[X^2] - E[X]^2$$

$$\sup |\psi''(\lambda)| = \sup |E[X^2] - E[X]^2|$$

$$= \sup |E[X_1^2, X_2^2, \dots, X_{k-1}^2, X] - E[X_1, X_2, \dots, X_{k-1}]|^2$$

c)  $X \sim N(\mu, (\frac{b-a}{2})^2)$

(Chebyshev's  
Upper bound)  $E[e^{\lambda X}] = e^{\mu\lambda + \lambda^2(\frac{b-a}{2})^2}$

$$\inf [\log E[e^{\lambda(X-\mu)}] - \lambda t] = \inf \left\{ \frac{\lambda^2(\frac{b-a}{2})^2}{2} - \lambda t \right\} \leq \frac{2t^2}{(b-a)^2}$$

2.5.

a)  $E[e^{\lambda X}] \leq e^{\frac{\lambda^2 \sigma^2}{2} + \lambda \mu}$  for all  $\lambda \in \mathbb{R}$

The derivation comes from a Taylor Expansion:

$$E[e^{\lambda X}] = 1 + \sum_{k=1}^n \frac{\lambda^k E[X^k]}{k!}$$

$$= 1 + \frac{\lambda E[X]}{1!} + \frac{\lambda^2 E[X^2]}{2!} + \dots$$

$$= 1 + \lambda \mu + \frac{\lambda^2 \sigma^2}{2} + \dots$$

So,  $E[X] = \mu$

b) From part a),  $E[X^2] = \sigma^2$

c. Yes,  $\text{Var}(X) = \sigma^2$  is also the lowest value standard deviation possible. The inequality arises from the removal of "1" - the first term.

(One Sided Bernstein's inequality)

$$\mathbb{E}[e^{\lambda(X-\mathbb{E}[X])}] \leq \exp\left(\frac{\lambda^2 \mathbb{E}[X^2]}{1 - \frac{\lambda b}{3}}\right), \text{ if } \lambda \in [0, 3/b]$$

$$\begin{aligned}
 2.6. P[Z_n \leq \mathbb{E}[Z_n] - \sigma^2 \delta] &= P[e^{\delta(Z - \mathbb{E}[Z])} \leq e^{-\sigma^2 \delta}] \\
 &\leq P\left\{ e^{\delta(1 + \mathbb{E}[z] + \frac{\delta \mathbb{E}[z^2]}{2} - \mathbb{E}[z])} \leq e^{-\sigma^2 \delta}\right\} \\
 &\leq P\left[ e^{\delta(1 + \mathbb{E}[z] + \frac{\delta \mathbb{E}[z^2]}{2} - \mathbb{E}[z]) h(\lambda b)} \leq e^{-\sigma^2 \delta}\right] \\
 &\leq P\left[ e^{\delta \frac{\delta^2 \mathbb{E}[z^2]}{2}} \left[\frac{1}{1 - \delta b / 3}\right] \leq e^{-\sigma^2 \delta}\right] \\
 P[Z_n \leq \mathbb{E}[Z] - \sigma^2 \delta] &\leq \exp\left(\frac{n \sigma^2 \delta^2}{2 \mathbb{E}[Z^2]}\right)
 \end{aligned}$$

$$\begin{aligned}
 \mathbb{E}[Z^2] = \mathbb{E}[X^4] &= \int_0^\infty [x^4] \cdot P(|X|) dx \\
 &= 8 \sigma^2
 \end{aligned}$$

$$P[Z_n \leq \mathbb{E}[Z] - \sigma^2 \delta] \leq \exp\left(-\frac{n \sigma^2 \delta^2}{16}\right)$$

$$\begin{aligned}
 2.7. a. \log \mathbb{E}[e^{\lambda X_i}] &= \left[ \sum_{k=0}^{\infty} \frac{\lambda^k \mathbb{E}[X]^k}{k!} \right] = \left[ 1 + \sum_{k=2}^{\infty} \frac{\lambda^k \mathbb{E}[X^2] \dots \mathbb{E}[X^{k-2}]}{k!} \right]
 \end{aligned}$$

$$\begin{aligned}
 &= \left[ 1 + \sum_{k=2}^{\infty} \frac{\lambda^k \sigma^2 \dots b^{k-2}}{k!} \right] = \left[ 1 + \frac{\sigma^2}{b^2} \sum_{k=2}^{\infty} \frac{\lambda^k b^k}{k!} \right] \\
 &= \left[ 1 + \frac{\sigma^2}{b^2} [e^{\lambda b} - 1 - \lambda b] \right]
 \end{aligned}$$

$$\log \mathbb{E}[e^{\lambda X_i}] \leq \frac{\sigma^2}{b^2} \left\{ \frac{e^{\lambda b} - 1 - \lambda b}{(b^2)} \right\}$$

$$\begin{aligned}
 b. P[\sum X_i \geq n \delta] &\leq e^{-n \sigma^2 \left\{ \frac{e^{\lambda b} - 1 - \lambda b}{b^2} \right\} - n \delta} \\
 &\leq \arg \min_{\lambda} \left\{ n \sigma^2 \left( \frac{e^{\lambda b} - 1 - \lambda b}{b^2} \right) - n \delta \right\}
 \end{aligned}$$

$$= n \frac{\sigma^2}{b} (e^{\lambda b} - 1) - n \bar{J}$$

$$\lambda^* = 0$$

$$\lambda = \frac{1}{b} \log \left( 1 + \frac{b\bar{J}}{\sigma^2} \right)$$

An introduction of maximum probability coefficient:

$$\begin{aligned} P[\sum \chi_i \geq t] &\leq e^{-\frac{n\sigma^2}{b^2} \{ e^{\lambda b} - 1 - \lambda b \bar{J} - n \bar{J} \}} \\ &\leq e^{-\frac{n\sigma^2}{b^2} \left[ \left( 1 + \frac{b\bar{J}}{\sigma^2} \right) - 1 - \log \left( 1 + \frac{b\bar{J}}{\sigma^2} \right) \right]} - \frac{n}{b} \log \left( 1 + \frac{b\bar{J}}{\sigma^2} \right) \\ &\leq e^{-\frac{n\sigma^2}{b^2} \left[ \left( 1 + \frac{b\bar{J}}{\sigma^2} \right) \log \left( 1 + \frac{b\bar{J}}{\sigma^2} \right) - \frac{b\bar{J}}{\sigma^2} \right]} \end{aligned}$$

$$h(t) = (1+t) \log(1+t) - t \quad \text{for } t \geq 0$$

$$\leq e^{-\frac{n\sigma^2}{b^2} \left[ h \left( \frac{b\bar{J}}{\sigma^2} \right) \right]}$$

C. Through a plot, Bernstein's  $\approx$  Bennett's

$$\frac{e^{\lambda b} - 1 - \lambda b}{b^2} \approx \frac{\lambda^2}{2(1 - b\lambda/3)}$$

2.8.

$$\text{a) } P[Z \geq t] \leq C e^{-\frac{t^2}{2(v^2 + b t)}}$$

$$\begin{aligned} E[Z] &= \int_1^\infty P[X \geq t] dt = \int_1^\infty P[X \geq t] dt + \int_2^\infty P[X \geq t] dt \\ &\leq \int_1^\infty C e^{-\frac{t^2}{4v^2}} dt + \int_2^\infty C e^{-\frac{t^2}{4bt}} dt \end{aligned}$$

Integral #1:

$$\int_1^\infty e^{-\frac{t^2}{4v^2}} dt = \int_0^{2\sqrt{\log C}} db + \int_{2\sqrt{\log C}}^\infty e^{-\frac{t^2}{4v^2}} dt$$

$$= 2\sqrt{\log C} + \sqrt{\pi} (|v| - \operatorname{erf}(\sqrt{\log C}))$$

$$= 2v(\sqrt{\pi} + \sqrt{\log C})$$

## Integral #2:

$$\int_2^{-t^2/4b} e^{-dt} = \int_0^{4b \log c} dt + \int_{4b \log c}^{\infty} e^{-t/4b} dt$$

$$= 4b \log c + 4b \cdot c \quad \text{where } c > 1.$$

$$\leq 4b(1 + \log c)$$

$$\mathbb{E}[Z] \leq 2v(\sqrt{\pi} + \sqrt{\log c}) + 4b(1 + \log c)$$

(Bernstains Condition)

$$[\mathbb{E}[(X-\mu)_+^k]] \leq \frac{1}{2} k! \sigma^2 b^{k-2} \text{ for } k=2,3,4$$

$$\begin{aligned} b) \mathbb{E}\left[\left|\frac{1}{n} \sum_{i=1}^n X_i\right|\right] &= \int P[X \geq t] dt = \int_1^{\infty} P[X \geq t] dt + \int_2^{\infty} P[X \geq t] dt \\ &= \int_0^{\frac{2\sigma \log c}{\sqrt{n}}} dt + \int_{\frac{2\sigma \log c}{\sqrt{n}}}^{\infty} e^{-\frac{t^2}{4\sigma^2}} dt + \int_0^{\frac{4b \log c}{n}} dt + \int_{\frac{4b \log c}{n}}^{\infty} e^{-\frac{t}{4b}} dt \end{aligned}$$

$$\leq \frac{2\sigma}{\sqrt{n}} (\sqrt{\pi} + \sqrt{\log 2}) + \frac{4b}{n} (1 + \log 2)$$

where  $c=2$  and  $v=\sigma$ .

(Kullback Leibler Divergence)

$$D(\delta || \alpha) = \delta \log\left(\frac{\delta}{\alpha}\right) + (1-\delta) \log\left(\frac{1-\delta}{1-\alpha}\right)$$

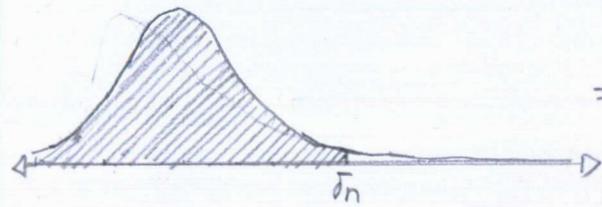
2.9.

$$\begin{aligned} a) P[Z_n \leq \delta_n] &:= P[e^{\lambda Z_n} \leq e^{\lambda \delta_n}] \\ &= P[\mathbb{E}[e^{\lambda Z_n}] \leq \mathbb{E}[e^{\lambda \delta_n}]] \end{aligned}$$

$$= P\left[\sum_n Z_n e^{\lambda n} \leq \mathbb{E}[e^{\lambda \delta_n}]\right]$$

$$= P\left[\sum_n \binom{n}{k} \alpha^k (1-\alpha)^{n-k} e^{\lambda n} \leq \mathbb{E}[e^{\lambda \delta_n}]\right]$$

Binomial Moment Generating Function



$$\begin{aligned}
 &= P\left[\sum_{k=1}^n \binom{n}{k} (\lambda e^\lambda)^k (1-\lambda)^{n-k} \leq \mathbb{E}[e^{\lambda \delta n}]\right] \\
 &= P\left[(1-\lambda + \lambda e^\lambda)^n \leq \mathbb{E}[e^{\lambda \delta n}]\right] \\
 &= \frac{(1-\lambda + \lambda e^\lambda)^n}{e^{\lambda \delta n}} \\
 \min_\lambda \left[ \frac{(1-\lambda + \lambda e^\lambda)^n}{e^{\lambda \delta n}} \right] = -\delta n + \frac{n \lambda e^\lambda}{(1-\lambda + \lambda e^\lambda)} = 0
 \end{aligned}$$

$$\lambda^* = \log \frac{\delta(1-\lambda)}{\lambda(1-\delta)} = \log \frac{(1-\lambda)}{(1-\delta)} - \log \left(\frac{\lambda}{\delta}\right)$$

$$\begin{aligned}
 P[Z_n \leq n\delta] &\leq e^{\{-n\delta\lambda^* - \log(1-\lambda - \lambda e^{\lambda^*})^n\}} \\
 &\leq e^{\{-n\left[\delta\left(\log\frac{(1-\lambda)}{(1-\delta)} - \log\frac{\lambda}{\delta}\right) - \log(1-\lambda - \lambda e^{\lambda^*})\right]\}} \\
 &\leq e^{-n\left[(\delta-1)\log\frac{(1-\lambda)}{(1-\delta)} - \delta\log\left(\frac{\lambda}{\delta}\right)\right]} \\
 &\leq e^{-nD(\delta||\lambda)} \\
 &\leq e
 \end{aligned}$$

(Hoeffding Bound)

$$P\left[\sum_{i=1}^n (X_i - \mu_i) \geq t\right] \leq \exp\left\{\frac{-t^2}{2\sum_{i=1}^n \sigma_i^2}\right\}$$

$$\begin{aligned}
 b) P[Z_n \leq \delta n] &= P[Z_n - \lambda n \leq \delta n - \lambda n] \\
 &\leq e^{(-n(\delta-\lambda)^2)} \quad (\text{Similarly, Question 2.3c}) \\
 &\leq e^{\left\{-\frac{t^2}{2\sum_{i=1}^n \sigma_i^2}\right\}}
 \end{aligned}$$

When  $\delta = \lambda$ , then the functions evaluate as one.

The second derivative of  $e^{-nD(\delta||\lambda)} \leq e^{-n(\delta-\lambda)^2} \leq e^{-t^2/2\sum_{i=1}^n \sigma_i^2}$

at  $\lambda = \delta$ .

$$2.11 \text{ a) } E[Z_n] = \int_0^\infty P[U > \delta] d\delta = \\ = \int_0^{\max} d\delta + n \int_{\max}^\infty P[U > x] dx$$

$$\operatorname{argmax}_x \left\{ n e^{-x^2/2\sigma^2} \right\} = 0$$

$$X^* = \sqrt{2\sigma^2 \log(n)}$$

$$E[Z_n] = \sqrt{2\sigma^2 \log(n)} + \int_{\sqrt{2\sigma^2 \log(n)}}^\infty P[U > x] dx$$

$$= \sqrt{2\sigma^2 \log(n)} + \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2\sigma^2 \log(n)}}^\infty e^{-x^2/2\sigma^2} dx$$

$$\leq \sqrt{2\sigma^2 \log(n)} + \left( \frac{4\sigma^2}{\sqrt{2\sigma^2 \log(n)}} \right) \sqrt{\frac{2}{\pi}} \int_{\sqrt{2\sigma^2 \log(n)}}^\infty e^{-x^2/2} dx$$

$$\leq \sqrt{2\sigma^2 \log(n)} + \sqrt{\frac{2}{\pi}} \sqrt{\frac{\pi}{2}} \frac{4\sigma^2}{\sqrt{2\sigma^2 \log(n)}} \operatorname{erf}\left(\frac{\delta}{\sqrt{2}}\right)$$

$$\leq \sqrt{2\sigma^2 \log(n)} + \frac{4\sigma^2}{\sqrt{2\sigma^2 \log(n)}}$$

$$\leq \sqrt{2\sigma^2 \log(n)} + \frac{4\sigma^2}{\sqrt{2\log(n)}}$$

Supposedly, the book's approximation is a Gaussian "lower tail bound".  
- Mills Ratio -

$$\text{b) } E[Z, n \geq 5] = E[\max |X|, n \geq 5]$$

$$= \int_0^\infty P[\max |X|, n \geq 5] dx$$

$$= \int_0^{\max} dx + \int_{\max}^\infty P[\max |X|, n \geq 5] dx$$

$$= \int_0^{\sqrt{2\sigma^2 \log(n)}} dx + \int_{\sqrt{2\sigma^2 \log(n)}}^\infty (1 - \phi(x))^n dx$$

Erf-Function:
$\prod_{i=1}^n e^{-x_i^2/2\sigma^2} = [e^{-x^2/2\sigma^2}]^n$
$= [1 - \phi(x)]^n$

$$= \sqrt{2\sigma^2 \log(n)} + \sqrt{2\sigma^2 \log(n)} \left(1 - \frac{1}{n}\right)^n$$

$$= (1 - \frac{1}{n}) \sqrt{2\sigma^2 \log(n)}$$

At large n-values,
$(1 - \frac{1}{n})^n = \frac{1}{e}$

2.10

$$a) \frac{1}{n} \log P[Z_n \leq \delta n] = \frac{1}{n} P[\log Z_n \leq \log \delta n]$$

$$= \frac{1}{n} P[\log \sum X_i \leq \log \delta n]$$

$$= \frac{1}{n} P[\log \sum \binom{n}{m} \alpha^m (1-\alpha)^{n-m} \leq \log \delta n]$$

$$= P\left[\frac{1}{n} \log \binom{n}{m} + \frac{m}{n} \log \alpha + \frac{(n-m)}{n} \log (1-\alpha) \leq \log \delta n\right]$$

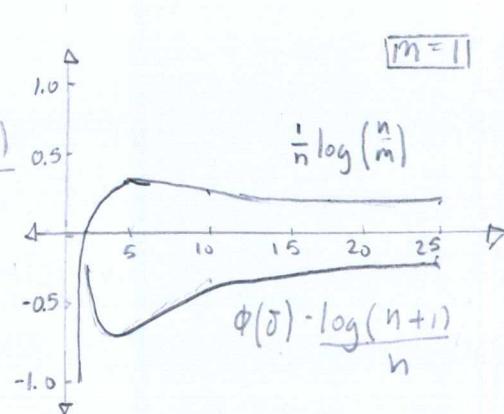
$$= P\left[\frac{1}{n} \log \binom{n}{m} + \bar{\delta} \log \alpha + (1-\bar{\delta}) \log (1-\alpha) \leq \log \delta n\right]$$

$$\leq \frac{1}{n} \log \binom{n}{m} + \bar{\delta} \log \alpha + (1-\bar{\delta}) \log (1-\alpha) - \log \delta n$$

$$b) \frac{1}{n} \log \binom{n}{m} \geq \phi(\bar{\delta}) - \frac{\log(n+1)}{n}$$

$$\textcircled{1} \text{ The } \lim_{n \rightarrow \infty} \frac{1}{n} \log \binom{n}{m} \geq \phi(\bar{\delta}) - \frac{\log(n+1)}{n}$$

\textcircled{2} By graph, the inequality remains true.



$$c) P[Z_n \leq \delta n] := P[e^{\lambda Z_n} \leq e^{\lambda \delta n}]$$

$$= P[\mathbb{E}[e^{\lambda Z_n}] \leq \mathbb{E}[e^{\lambda \delta n}]]$$

$$= P\left[\sum_k \mathbb{E}[e^{\lambda k}] \leq \mathbb{E}[e^{\lambda \delta n}]\right]$$

$$= P\left[\sum_k \binom{n}{k} \alpha^k (1-\alpha)^{n-k} e^{\lambda k} \leq \mathbb{E}[e^{\lambda \delta n}]\right]$$

$$= P\left[\sum_k \binom{n}{k} (\alpha e^\lambda)^k (1-\alpha)^{n-k} \leq \mathbb{E}[e^{\lambda \delta n}]\right]$$

$$= P[(1-\alpha + \alpha e^\lambda)^n \leq \mathbb{E}[e^{\lambda \delta n}]]$$

$$\underset{\lambda}{\operatorname{argmin}} \left[ (1-\alpha + \alpha e^\lambda)^n / \mathbb{E}[e^{\lambda \delta n}] \right] = 0$$

$$\lambda^* = \log \frac{\sigma(1-\alpha)}{\alpha(1-\delta)} = \log \frac{(1-\alpha)}{(1-\delta)} - \log \left( \frac{\alpha}{\delta} \right)$$

$$\begin{aligned} P[e^{\lambda Z_n} \leq e^{\lambda \delta n}] &= e^{-n \delta \lambda^* - \log(1-\alpha - \alpha e^{\lambda^*})^n} \\ &= e^{-n[(\delta-1)\log(\frac{1-\alpha}{1-\delta}) - \delta \log(\frac{\alpha}{\delta})]} \\ &= e^{-n D(\delta || \alpha)} \geq e^{-n D(\delta || \alpha)} \end{aligned}$$

$$2.12. \text{ a) } \mathbb{E}[\max_{i=1 \dots n} X_i] = \sum \mathbb{E}[e^{\lambda X_i}] \leq n e^{-\lambda^2 \sigma^2 / 2} = 0$$

$$\log \mathbb{E}[\max e^{\lambda X_i}] \leq \log n - \frac{\lambda^2 \sigma^2}{2} \approx 0 \quad \text{because an upper tail bound}$$

$$\lambda^* = \sqrt{2 \log n} / \sigma$$

$$\mathbb{E}[\max X_i] \leq \log n - \lambda^2 \left( \frac{\sigma^2}{2} \right)$$

$$\frac{\mathbb{E}[\max X_i]}{\lambda} \leq \frac{\log n}{\lambda} - \lambda^2 \left( \frac{\sigma^2}{2} \right)$$

Authors across the pond scale  $\lambda$  at the beginning.

$$\begin{aligned} b) \quad & \frac{\mathbb{E}[\max X_i]}{\lambda} \leq \frac{\sigma \sqrt{\log n}}{\sqrt{2}} - \frac{\sigma \sqrt{\log n}}{\sqrt{2}} \\ & \leq \sqrt{2 \sigma^2 \log n} \end{aligned}$$

$$b) \quad \frac{\mathbb{E}[\max X_i]}{\lambda} \leq \sqrt{2 \sigma^2 \log n}$$

$$\mathbb{E}[\max X_i] \leq 2 \sqrt{\log n} \leq 2 \sqrt{\sigma^2 \log n}$$

$$2.13 a) P[X_1 + X_2] = \mathbb{E}[e^{\lambda(X_1 + X_2)}]$$

$$= \mathbb{E}[e^{\lambda[(1+\mathbb{E}[X_1] + \lambda \mathbb{E}[X_1^2]) + (1+\mathbb{E}[X_2] + \lambda \mathbb{E}[X_2^2])]}]$$

$$\leq \mathbb{E}[e^{\lambda^2 (\mathbb{E}[X_1^2] + \mathbb{E}[X_2^2]) / 2}]$$

$$\leq \mathbb{E}[e^{\lambda^2 (\sigma_1^2 + \sigma_2^2) / 2}]$$

$$\leq e^{\lambda^2 (\sqrt{\sigma_1^2 + \sigma_2^2}) / \sqrt{2}}$$

$$b) P[X_1 + X_2] = \mathbb{E}[e^{\lambda(X_1 + X_2)}]$$

$$\leq e^{\lambda(\sigma_1^2 + \sigma_2^2) / 2}$$

$$\leq e^{\lambda \sqrt{2} \sqrt{\sigma_1^2 + \sigma_2^2} / 2}$$

$$c) P[X_1 + X_2] = \mathbb{E}[e^{\lambda(X_1 + X_2)}]$$

$$\leq e^{\lambda(\sigma_1^2 + \sigma_2^2) / 2}$$

$$d) P[X_1 \circ X_2] = \mathbb{E}[e^{\lambda(X_1 \circ X_2)}]$$

$$\begin{aligned}
 &= \mathbb{E}[e^{(1+\lambda\mathbb{E}[X]+\lambda^2\mathbb{E}[X^2]/2)(1+\lambda\mathbb{E}[X_2]+\lambda^2\mathbb{E}[X_2^2]/2)}] \\
 &\leq \mathbb{E}\left[e^{\frac{\lambda^2\mathbb{E}[X^2]}{2}}\right]\left[\frac{\lambda^2\mathbb{E}[X_2^2]}{2}\right] \\
 &\leq \mathbb{E}[e^{\lambda^4\sigma_1^2\sigma_2^2/4}] \quad (\text{V, b parameters: } e^{\lambda^2/2})
 \end{aligned}$$

$$(V, b) = (\sqrt{2}\sigma_1\sigma_2, \sqrt{2}\sigma_1\sigma_2)$$

2.14.

$$\begin{aligned}
 a) P[X - \mathbb{E}[X] \geq t] &:= P[e^{\lambda(X-\mathbb{E}[X])} \geq e^{\lambda t}] \\
 &= \mathbb{E}[e^{\lambda(1+\mathbb{E}[X]+\lambda\mathbb{E}[X^2]/2)-\lambda\mathbb{E}[X]} \geq e^{\lambda t}] \\
 &\leq \mathbb{E}[e^{\lambda^2\mathbb{E}[X^2]/2} \geq e^{\lambda t}] \\
 &\leq e^{\lambda^2\sigma_2^2/2 - \lambda t}
 \end{aligned}$$

$$\operatorname{argmax}_{\lambda} \left\{ \lambda^2\sigma_2^2/2 - \lambda t \right\} = 0$$

$$\lambda^* = \frac{t}{\sigma^2}$$

$$\begin{aligned}
 P[X - \mathbb{E}[X] \geq t] &\leq e^{\lambda^2\sigma_2^2/2 - \lambda t} \\
 &\leq e^{-t^2/2\sigma^2}
 \end{aligned}$$

$$C_1 = 1; C_2 = \sqrt{2}\sigma^2; \operatorname{Var}(X) \leq \frac{C_1}{C_2}; \sigma \leq \sqrt{2}\sigma$$

$$b) \mathbb{E}[|X - m_X|] = \int_{-\infty}^{m_X} P[X \leq m_X] dx + \int_{m_X}^{\infty} P[X \geq m_X] dx$$

Uniqueness of  $X^*$

$$\mathbb{E}[|X - m_X|] = 0 = \int_{-\infty}^{m_X} P[X \leq m_X] dx + \int_{m_X}^{\infty} P[X \geq m_X] dx$$

$$\int_{-\infty}^{m_X} P[X \leq m_X] dx = \int_{m_X}^{\infty} P[X \geq m_X] dx$$

Uniqueness of another  $X$ :

$$\mathbb{E}[|X-m_X|] = c = \int_{-\infty}^{m_X} P[X \leq m_X] dx + \int_{m_X}^{\infty} P[X \geq m_X] dx$$

$$\int_{-\infty}^{m_X} P[X \leq m_X] dx = c + \int_{-\infty}^{m_X} P[X \leq m_X] dx$$

c)  $P[|X-m_X| \geq t] = P[|X-m_X|^2 \geq t^2]$

$$= P\left[\frac{c_2}{3}|X-m_X|^2 \geq \frac{c_2}{3}t^2\right]$$

$$\leq e^{-\frac{c_2}{3}t^2} \cdot \mathbb{E}\left[\frac{c_2}{3}|X-m_X|^2\right]$$

$$\leq e^{-\frac{c_2}{3}t^2} \cdot \int_{-\infty}^{\infty} x^2 e^{-\frac{c_2}{3}x^2} dx$$

$$\leq \frac{4}{c_2} e^{-\frac{c_2}{3}t^2}$$

$$\leq \frac{4c_1}{c_2} e^{-\frac{c_2}{3}t^2}$$

$$\leq \frac{c_3}{c_2} e^{-\frac{c_4}{4}t^2} \quad \text{where } c_3 = 4c_1 \text{ and } c_4 = \frac{c_2}{8}$$

$$\leq c_3 e^{-\frac{c_4}{4}t^2}$$

d)

$$\leq c_3 e^{-\frac{c_4}{4}t^2}$$

d)  $P[|X-m_X| \geq t] = P[|X-m_X|^2 \geq t^2]$

$$= P\left[\frac{c_4}{4}|X-m_X|^2 \geq \frac{c_4}{4}t^2\right]$$

$$= P\left[\frac{c_4}{4}|X-E[X]|^2 \geq \frac{c_4}{4}\left(\frac{t}{2}\right)^2 + \frac{c_4}{4}\left(\frac{t}{2}\right)^2\right]$$

$$|X - E[X]| = |X - m_X| + \frac{c_4}{4}\left(\frac{t}{2}\right)$$

$$P[|X-m_X|^2 \geq t^2] = P\left[\frac{c_4}{4}|X-E[X]|^2 \geq \frac{c_4}{4}\left(\frac{t}{2}\right)^2\right]$$

$$= e^{-\frac{c_4}{4}\left(\frac{t}{2}\right)^2} \cdot \int_{\frac{t}{2}}^{\infty} x e^{-\frac{c_4}{4}x^2} dx$$

$$\begin{aligned}
 &\leq \frac{C_1}{C_4} e^{-\frac{C_4}{8} t^2} \text{ where } C_2 = \frac{C_4}{8} \text{ and } C_1 = 2C_3 \\
 &\leq \frac{C_3}{C_4} e^{-\frac{C_2 t^2}{2}} \\
 &\leq C_3 e^{-\frac{C_2 t^2}{2}} \quad \text{when } C_4 \geq 1
 \end{aligned}$$

2.15. a)  $\hat{F}_n(x) := \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) ; \int_{-\infty}^{\infty} K(t) dt = 1 ; h > 0$

$$\begin{aligned}
 P[\|\hat{F}_n - F\|] &= \frac{1}{nh} \int_{-\infty}^{\infty} K\left(\frac{x-\hat{X}_i}{h}\right) - K\left(\frac{x-X_i}{h}\right) dx \\
 &= \frac{2}{n} \int_{-\infty}^{\infty} K\left(\frac{x-X_i}{h}\right) dx \\
 &= \frac{2}{n}
 \end{aligned}$$

$$\begin{aligned}
 P[\|\hat{F}_n - F\| \geq \mathbb{E}[\|\hat{F}_n - F\|] + \delta] &= P[\|\hat{F}_n - F\| - \mathbb{E}[\|\hat{F}_n - F\|] \geq \delta] \\
 &= P[e^{\lambda(\|\hat{F}_n - F\| - \mathbb{E}[\|\hat{F}_n - F\|])} \geq e^{\lambda \delta}] \\
 &\leq e^{-\frac{\lambda^2 \sigma^2}{n}} / e^{\lambda \delta}
 \end{aligned}$$

$$\arg \min_{\lambda} \left\{ \frac{\lambda^2 \sigma^2}{n} - \lambda \delta \right\} = 0$$

$$\lambda^* = \frac{\delta n}{2\sigma^2}$$

$$\left( \frac{\delta n}{2\sigma^2} \right)^2 \left( \frac{\sigma^2}{n} \right)$$

$$\begin{aligned}
 P[\|\hat{F}_n - F\| - \mathbb{E}[\|\hat{F}_n - F\|] \geq \delta] &\leq e^{-\frac{n \delta^2}{4}} / e^{\frac{\delta^2}{2}}
 \end{aligned}$$

$$\leq e^{-n \delta^2 / 8}$$

(Hilbert Space)

$$\forall \langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$$

$$1) \langle f, g \rangle_V = \langle g, f \rangle_V$$

for all  $f, g \in V$

$$2) \langle f, f \rangle_V \geq 0$$

for all  $f \in V$ , with equality iff  $f = 0$

$$3) \langle f + \alpha g, h \rangle_V = \langle f, h \rangle_V + \alpha \langle g, h \rangle_V \text{ for all } f, g, h \in V \text{ and } \alpha \in \mathbb{R}$$

Also, Hilbert Space ( $\mathbb{H}$ ) is an inner product  $(\langle \cdot, \cdot \rangle_{\mathbb{H}}, \|\cdot\|_{\mathbb{H}})$   
 in which every Cauchy sequence  $(f_n)_{n=1}^{\infty}$  in  $\mathbb{H}$  converges to  $f^* \in \mathbb{H}$ .

2.16.

$$a) \delta \geq 0 ; S_n = \left\| \sum X_i \right\|_{\mathbb{H}} ; b^2 = \frac{1}{n} \sum_{i=1}^n \|X_i\|_{\mathbb{H}}^2 ; \left\| X_i \right\|_{\mathbb{H}} \leq b_i$$

$$P[|S - E[S_n]| \geq n\delta] := E[e^{\lambda(S - E[S_n])}] \geq e^{\lambda n\delta}$$

$$= E[e^{\lambda(1 + E[S_n] + \lambda E[S_n^2]/2 - \lambda E[S_n])}] \geq e^{\lambda n\delta}$$

$$\leq E[e^{\lambda^2 E[S_n^2]/2}] \geq e^{\lambda n\delta}$$

$$\leq E[e^{\lambda^2 b^2/2}] \geq e^{\lambda n\delta}$$

$$\arg \min_{\lambda} \left\{ \lambda^2 b^2/2 - \lambda n\delta \right\} = 0$$

$$\lambda^* = \frac{n\delta}{b^2}$$

$$\lambda^2 b^2/2$$

$$P[|S - E[S_n]| \geq n\delta] \leq \frac{e^{-\lambda n\delta}}{e^{-\lambda n\delta}}$$

$$\leq e^{-\frac{n\delta}{2b^2}}$$

$$\leq 2^{\circ} e^{-\frac{n\delta}{3b^2}}$$

$$b) P\left[\frac{S_n}{n} \geq \alpha + \delta\right] := P\left[e^{\frac{\lambda S_n}{n}} \geq e^{\lambda(\alpha + \delta)}\right]$$

$$= P\left[e^{\lambda(1 + E[\frac{S_n}{n}] + \lambda E[\frac{S_n^2}{n^2}]/2)} \geq e^{\lambda(\alpha + \delta)}\right]$$

$$\text{If } a = \sqrt{\frac{1}{n^2} \sum_{\mathbb{H}} E[\|X\|^2]}$$

$$P\left[\frac{S_n}{n} - a \geq \delta\right] \leq P\left[e^{\lambda^2 E[S_n^2/n]/2} \geq e^{\lambda\delta}\right]$$

$$\arg \min_{\lambda} \left\{ \lambda^2 E[S_n^2/n]/2 - \lambda\delta \right\} = 0$$

$$\lambda^* = \delta / E[S_n^2/n]$$

$$\leq \frac{e^{\left(\frac{n\delta}{\mathbb{E}[S_n^2]}\right)^2 \frac{\mathbb{E}[S_n^2]}{2}}}{e^{\left(\frac{n\delta}{\mathbb{E}[S_n^2]}\right) \cdot \delta}} \\ \leq e^{-\frac{n\delta^2}{2b^2}} \\ \leq e^{-\frac{n\delta^2}{8b^2}}$$

(Hanson-Wright Inequality) 2.17  
 $P[|X^TAX - \mathbb{E}[X^TAX]| > t] \leq 2 \cdot e^{[-c \cdot \min\left(\frac{t^2}{K^4 \|A\|_F^2}, \frac{t}{K^2 \|A\|_2}\right)]}$

where  $c > 0$ ,  $\|X\| \leq K$ ,  $\|A\|_F^2 = \sqrt{\sum A_{ii}^2}$ ,  $\|A\|_2 = \sup \|Ax\|_2$

2.17.

$$Z = \sum_{i=1}^n \sum_{j=1}^n A_{ij} X_i X_j$$

$$P[|X^TAX - \mathbb{E}[X^TAX]| > t] = P[|Z - \mathbb{E}[Z]| \geq t]$$

$$\leq P[e^{\lambda^2 \mathbb{E}[Z^2]/2} \geq e^{\lambda t}]$$

$$\leq e^{-\frac{t^2}{2\sigma^2}}$$

$$(v, b) \text{ parameters : } e^{-\frac{\lambda v^2}{b}} ; (v, b) = (t, 2\sigma^2 \lambda) \\ -\min\left(\frac{\lambda^2 \sigma^2}{2} - \lambda t\right)$$

$$P[|X^TAX - \mathbb{E}[X^TAX]| > t] = e^{-\min\left(\left(\frac{v}{b}\right)^2, \left(\frac{v}{b}\right)\right)} \\ = e^{-\min\left(\frac{t^2}{4\sigma^2 \lambda^2 \cdot \sigma^4}, \frac{t}{2\sigma^2 \lambda}\right)} \\ = e^{-\min\left(\frac{t^2}{K^4 \|A\|_F^2}, \frac{t}{K^2 \|A\|_2}\right)} \\ \leq 2 \cdot e^{-c \cdot \min\left(\frac{t^2}{K^4 \|A\|_F^2}, \frac{t}{K^2 \|A\|_2}\right)}$$

(Orlicz Norm) 2.18

$$\|X\|_q = \inf \{t > 0 \mid \mathbb{E}[q(E|X|)] \leq 1\}, \quad q(0) = 0$$

2.18.

a)  $P[|X| > t]$

$$P[|X|^q > t^q] = P[e^{\lambda |X|^q} > e^{\lambda t^q}]$$

$$= P\left[e^{\lambda(1+\mathbb{E}[X]+\lambda\mathbb{E}[X^2]/2)^2} > e^{\lambda t^2}\right]$$

$$\leq \frac{e^{\lambda^2\mathbb{E}[X^2]}}{e^{\lambda t^2}}$$

$$\operatorname{argmin}\{\lambda^2\mathbb{E}[X^2]/2 - \lambda t^2\} = 0$$

$$\lambda^* = \frac{t^2}{\mathbb{E}[X^2]^2} = \frac{t^2}{\|X\|^2}$$

$$P[|X|^2 > t^2] \leq e^{-\frac{t^2}{2\|X\|^2}}$$

$$\leq 2^0 e^{-\frac{t^2}{2\|X\|^2}}$$

$$\leq C_1 e^{-C_2 t^2}$$

where  $C_1 = 2$  and  $C_2 = \|X\|^2$

b) If  $1 \geq C_1 e^{-C_2 t^2}$

$$e^{\frac{C_2 t^2}{C_1}} - 1 \geq 0$$

$$\mathbb{E}[e^{\frac{C_2 t^2}{C_1}} - 1] = \int_0^\infty P(e^{\frac{C_2 t^2}{C_1}} - 1 > u) du$$

$$\leq \int_0^\infty P(C_2 > t^2 \log(1+u)) du$$

$$\leq C_1 \int_0^\infty e^{-C_2 t^2 \log(1+u)} du$$

$$\leq C_1 \int_0^\infty e^{-C_2 t^2 v} (1+v) dv \quad \text{when } v = \log(1+u)$$

$$= \frac{C_1(1-C_2 t^2)}{(1+C_2 t^2)^2}$$

$$\lim_{t \rightarrow \infty} \frac{C_1(1-C_2 t^2)}{C_2 t^2} = 0 \text{ and less than infinity.}$$

2.19. An example about inverse functions:

$$4 \left( \frac{\mathbb{E}[\max_{i=1 \dots n} |X_i|]}{\sigma} \right) \leq \mathbb{E} \left[ 4 \left( \frac{\max_{i=1 \dots n} |X_i|}{\sigma} \right) \right] \\ \leq \sum_{i=1}^n \mathbb{E} \left[ \frac{4(|X_i|)}{\sigma} \right] \\ \leq n$$

A function and its inverse both strictly increase when positive.

$$4 \left( \frac{\mathbb{E}[\max_{i=1 \dots n} |X_i|]}{\sigma} \right) \leq n$$

$$\mathbb{E} \left[ \frac{\max_{i=1 \dots n} |X_i|}{\sigma} \right] \leq 4^{-1}(n)$$

$$\mathbb{E}[\max_{i=1 \dots n} |X_i|] \leq \sigma 4^{-1}(n)$$

(Rosenthal's Inequality)

$$\mathbb{E} \left[ \left( \sum_{i=1}^n X_i \right)^{2m} \right] \leq R_m \left\{ \sum_{i=1}^n \mathbb{E}[X_i^{2m}] + \left( \sum_{i=1}^n \mathbb{E}[X_i^2] \right)^{2m} \right\}$$

$$2.20. \|X_i\|_{2m} = (\mathbb{E}[X_i^{2m}])^{1/2m} \leq C_m$$

$$P \left[ \left| \frac{1}{n} \sum_{i=1}^n X_i \right| \geq \delta \right]$$

$$P \left[ \left| \frac{1}{n} \sum_{i=1}^n X_i \right|^{2m} \geq \delta^{2m} \right] := P \left[ e^{\lambda \left( \frac{1}{n} \sum_{i=1}^n X_i \right)^{2m}} \geq e^{\lambda \delta^{2m}} \right]$$

$$= P \left[ e^{\lambda (1 + \mathbb{E}[X_i^{2m}] + \lambda \mathbb{E}[X_i^{2m}]^2 / n^{2m})} \geq e^{\lambda \delta^{2m}} \right]$$

$$\leq P \left[ e^{\frac{\lambda^2 \mathbb{E}[X_i^2]^{2m}}{2n^{2m}}} \geq e^{\lambda \delta^{2m}} \right]$$

$$\arg \min_{\lambda} \left\{ \frac{\lambda^2 \mathbb{E}[X_i^2]^{2m}}{2n^{2m}} - \lambda \delta^{2m} \right\} = 0$$

$$\lambda^* = \left( \frac{n \delta^n}{\mathbb{E}[X_i^2]} \right)^{2m} \leq \left( \frac{\delta^n}{C_m} \right)^{2m}$$

$$P[Z \geq z^{\delta}] \leq 2^{n(R - D(\delta || \kappa))} \quad \text{where } N = 2^{nR}$$

b) i)  $R - D(\delta || \kappa) > 0$

$$P[Z^{\delta} \geq z^{\delta}] = 2^{n(R - D(\delta || \kappa))}$$

$$\geq \frac{\mathbb{E}[V]^2}{\mathbb{E}[V^2]} \quad \text{where Bernoulli mean } \kappa \quad \text{variance } = \kappa(1-\kappa)$$

$$\geq \frac{\kappa^2}{\kappa(1-\kappa)} \quad \text{at } \kappa = 1/2$$

$$\geq 1 \quad n(R - D(\delta || \kappa))$$

$$\text{ii) } \lim_{n \rightarrow \infty} P[Z^{\delta} \geq z^{\delta}] = \lim_{n \rightarrow \infty} 2^{-n(R - D(\delta || \kappa))} = 0$$

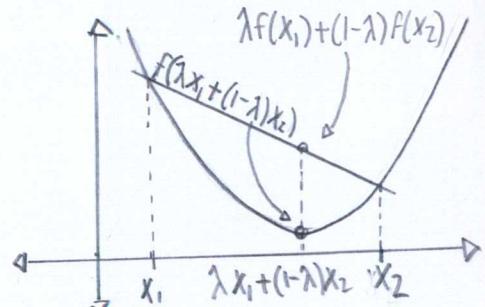
2.22.

$$P_{\theta}(x_1, \dots, x_d) = \exp \left\{ \frac{1}{\sqrt{d}} \sum_{i,j} \theta_{ijk} x_i x_k - F_d(\theta) \right\}$$

$$F_d(\theta) = \log \left( \sum_{i=1}^n \exp \left\{ \frac{1}{\sqrt{d}} \sum_{j,k} \theta_{ijk} x_j x_k \right\} \right)$$

(convex Function).

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$



$$\text{a) } F_d(\theta) = \log \left( \sum_{i=1}^n \exp \left\{ \frac{1}{\sqrt{d}} \sum_{j,k} \theta_{ijk} x_j x_k \right\} \right)$$

$$= \log \left( \sum_{i=1}^n \exp \left\{ \frac{1}{\sqrt{d}} [\theta x_i + (1-\theta)x_2] \right\} \right)$$

$$\leq \theta \log \left( \sum_{i=1}^n \exp \left\{ \frac{x_1}{\sqrt{d}} \right\} \right) + (1-\theta) \log \left( \sum_{i=1}^n \exp \left\{ \frac{x_2}{\sqrt{d}} \right\} \right)$$

$$\text{b) } \frac{\|F_d(\theta) - F_d(\theta')\|}{\|\theta - \theta'\|} = \|\nabla_{\theta} F_d(\theta)\|$$

$$= \frac{1}{\sqrt{d}} \sum_{j,k} x_j x_k \quad \text{where } |x_j x_k| = 1$$

$$\leq \sqrt{d} \quad \text{for the set } \{1, 2, \dots, d\}$$

$$\text{So, } \|F_d(\theta) - F_d(\theta')\| \leq \sqrt{d} \|\theta - \theta'\|$$

$$\begin{aligned}
 P\left[\left|\frac{1}{n} \sum X_i\right|^{2m} \geq \delta^{2m}\right] &\leq e^{\left(\frac{n\delta}{cm}\right)^{2m} \alpha \left(\frac{cm}{2n}\right)^{2m} - \left(\frac{n\delta}{cm}\right)^{2m} \delta^{2m}} \\
 &\leq e^{-\left(\frac{n\delta}{2cm}\right)^{2m}} \quad cm = n \frac{3/2 \delta^3}{2} \\
 &\leq e^{-\left(\frac{1}{\sqrt{n}\delta}\right)^{2m}} \\
 &\leq \beta_m \left(\frac{1}{\sqrt{n}\delta}\right)^{2m}
 \end{aligned}$$

(Rescaled Hamming Distortion)

$$d(X) = \min_{j=1..N} \rho_H(X, z^j) = \min_{j=1..N} \left\{ \frac{1}{n} \sum_{i=1}^n [X_i \neq z_i^j] \right\}$$

$$2.21. \text{ a)} R < D_2(\delta || 1/2) = \delta \log_2 \frac{\delta}{1/2} + (1-\delta) \log_2 \frac{1-\delta}{1/2}$$

$$\text{IF } V = \sum_{j=1}^n V_j, \text{ then } d(X) = \min_{j=1..N} \left\{ \frac{1}{n} V \right\}$$

$$\begin{aligned}
 P[d(X) \leq \delta] &= P\left[\min_{j=1..N} \frac{1}{n} V \leq \delta\right]; P\left[\min_{j=1..N} V \leq n\delta\right] \\
 &= N \cdot P[\lambda V \leq \lambda n \delta]
 \end{aligned}$$

$$\begin{aligned}
 N \cdot P[2^{\lambda V} \leq 2^{\lambda n \delta}] &= N \cdot P[\mathbb{E}[2^{\lambda V}] \leq \mathbb{E}[2^{\lambda n \delta}]] \\
 &= N \cdot P\left[\sum_k^\infty k 2^{\lambda k} \leq \mathbb{E}[2^{\lambda n \delta}]\right] \\
 &= N \cdot P\left[\sum_k^\infty \binom{n}{k} \alpha^k (1-\alpha)^{n-k} 2^{\lambda k} \leq \mathbb{E}[2^{\lambda n \delta}]\right] \\
 &= N \cdot P\left[\sum_k^\infty \binom{n}{k} (\alpha 2^\lambda)^k (1-\alpha)^{n-k} \leq \mathbb{E}[2^{\lambda n \delta}]\right] \\
 &= N \cdot P[(1-\alpha + \alpha 2^\lambda)^n \leq \mathbb{E}[2^{\lambda n \delta}]]
 \end{aligned}$$

$$\arg \min_{\lambda} \left\{ (1-\alpha + \alpha 2^\lambda)^n - \lambda n \delta \right\} = 0$$

$$\lambda^* = \log \frac{\delta(1-\alpha)}{\alpha(1-\delta)} = \log \frac{(1-\alpha)}{(1-\delta)} - \log \left(\frac{\alpha}{\delta}\right)$$

$$\begin{aligned}
 N \cdot P[2^{\lambda V} \leq 2^{\lambda n \delta}] &= N \cdot Z \\
 &= N \cdot Z^{-n D(\delta || \alpha)}
 \end{aligned}$$

$$c) \theta_{jk} \sim N(0, \beta^2)$$

$$\text{From the hint: } E[F_d(\theta)] = \lim_{d \rightarrow \infty} \frac{F_d(\theta)}{d} = \log 2 + \frac{\beta^2}{4}$$

$$P\left[\frac{F_d(\theta)}{d} \geq \log 2 + \frac{\beta^2}{4} + t\right] = P\left[\frac{F_d(\theta)}{d} \geq E[F_d(\theta)] + t\right]$$

$$= P\left[\lambda \frac{F_d(\theta)}{d} \geq \lambda E[F_d(\theta)] + \lambda t\right]$$

$$P\left[e^{\lambda\left(\frac{F_d(\theta)}{d} - E[F_d(\theta)]\right)} \geq e^{\lambda t}\right] = P\left[e^{\lambda\left(1 + E[F_d(\theta)] + \lambda E\left[\frac{F_d(\theta)^2}{d}\right] - E[F_d(\theta)]\right)} \geq e^{\lambda t}\right]$$

$$\leq P\left[e^{\frac{\lambda^2 \beta^2}{2d}} \geq e^{\lambda t}\right]$$

$$\leq \frac{e^{\frac{\lambda^2 \beta^2}{2d}}}{e^{\lambda t}}$$

$$\underset{\lambda}{\operatorname{argmin}} \left\{ \frac{\lambda^2 \beta^2}{2d} - \lambda t \right\} = 0$$

$$\lambda^* = \frac{dt}{\beta^2}$$

$$-\beta dt^2/2$$

$$P\left[e^{\lambda\left(\frac{F_d(\theta)}{d} - E[F_d(\theta)]\right)} \geq e^{\lambda t}\right] \leq e$$

$$\leq 2 \cdot e^{-\beta dt^2/2}$$

### Notes:

1) Symbol ( $\hat{=}$ ) implies a similar argument, although not truly equivalent.

2) Two like-methods from the chapter are:

$$A) P[z > t] = P[\lambda z \geq \lambda t]$$

$$P[e^{\lambda z} \geq e^{\lambda t}] \leq \frac{e^{\lambda z}}{e^{\lambda t}}$$

$$B) P[z > t] = P[\lambda z \geq \lambda t]$$

$$E[P[z > t]] = \frac{\int_t^\infty P[z > t] dz}{e^{\lambda t - \lambda t}}$$