Beweis. Unterscheiden zwei Faelle: Gilt $Q \ll P$, so gibt es Radon-Nikodym-Dichte $Y := \dfrac{dQ}{dP}$. Diese ist normiert $E(Y) = \int Y dP = \int 1 dQ = 1$. Bemerkungen 4.34 und 4.37 geben:

$$\begin{aligned} D(Q\|P) &= \sup_{0 \le T \in \mathcal{L}^1} E\left[Y(\log T - \log(ET))\right] \\ &= \sup_{0 \le T \in \mathcal{L}^1} \left[E_Q(\log T - (EY)\log(ET)\right] \\ &= \sup_{0 \le T \in \mathcal{L}^1} \left[E_Q(\log T - \log(ET)\right] \end{aligned}$$

wg. Normierung von Y. Nun führen wir Substitution $Z = \log T$, also $e^Z = T$ ein, so dass

$$E(e^Z) = E(T) < \infty$$

Also

$$D(Q\|P) = \sup_{Z \text{ ZV mit } e^Z \in \mathcal{L}^1} \left[E_Q(Z - \log(Ee^Z)\right]$$

Gilt $Q \ll P$ nicht, so gibt es ein Ereignis A mit $Q(A) > 0$ aber $P(A) = 0$ und nach Defn. gilt $D(Q\|P) = \infty$. Andererseits ergibt Einsetzten von $Z_n := n\mathbb{1}_A$ in r.S. (4.15)

$$\begin{aligned} &nE_Q(\mathbb{1}_A) - \log E\left[e^{n\mathbb{1}_A}\right] \\ =&nQ(A) - \log\left[\underbrace{P(A)}_{=0} e^n + P(A^c)\cdot 1\right] \\ =&nQ(A) - \log 1 = n\underset{\substack{\vee \\ 0}}{Q(A)} \to \infty \text{ falls } n \to \infty. \end{aligned}$$

Demnach sind beide Seiten in (4.15) $+\infty$. □

Bemerkung 4.40 (Verallgemeinerung von Bekanntem:). der Erwartungswert minimiert den mittleren quadratischen Abstand einer deterministishen Zahl zum zufaelligen Wert einer ZV. Dies ist ein Speziallfall von allgemeinerem Prinzip:

Satz 4.41. *Sei $I \subset \mathbb{R}$ offenes Intervall und $f\colon I \to \mathbb{R}$ konvex und diffbar. Sei $X\colon \Omega \to I$ eine ZV auf $(\Omega, \mathcal{A}, P)$. Dann*

$$E\left[f(X) - f(EX)\right] = \inf_{a \in I} E\left[f(X) - f(a) - f'(a)(X-a)\right]$$

Beweis. Buch/Praesenzuebung □

$\inf_{a\in I} g_a(x) \le \mathbb{E}\left[f(x) - f(\mathbb{E}X) - f'(\mathbb{E}x)(X - \mathbb{E}(x))\right] = \mathbb{E}\left[f(x) - f(\mathbb{E}x)\right]$

z.Z: $\inf_{a\in I} g_a(x) \ge \mathbb{E}\left[f(x) - f(\mathbb{E}x)\right]$

$$\inf_{a \in I} \mathbb{E}[\cancel{f(X)} - f(a) - f'(a)(X-a)] - \mathbb{E}[\cancel{f(X)} - f(\mathbb{E}X)] \geq 0$$

$$\inf_{a \in I} f(\mathbb{E}X) - f(a) - f'(a)(\mathbb{E}X - a) \geq 0$$ aus Konvexität ✓

Bemerkung 4.42 (Bregman-Divergenz). Fuer solches f wird fuer $x, y \in I$

$$g_x(y) := \left[f(y) - f(x) - f'(x)(y-x)\right]$$

die *Bregman-Divergenz* von x nach y genannt. Fuer Konvexe f ist sie nicht-negativ.

Wir fuehren den erwaehnten, bekannten Fall des EW auf den Satz zurueck: Sei $f(x) = x^2$, dann $f'(x) = x2$, $f'(a) = 2a$. Betrachte r.S.

$f(x) - f(a) - f'(a)(x-a) = x^2 - a^2 - 2a(x-a) = x^2 - a^2 - 2ax + 2a^2 = (x-a)^2$

Nun l.S.

$$\begin{aligned} E\left[f(X) - f(EX)\right] &= E\left[X^2 - (EX)^2\right] \\ &= E\left[X^2\right] - (EX)^2 \\ &= E\left[X^2\right] - 2(EX)^2 + (EX)^2 \\ &= E\left[X^2 - 2X(EX) + (EX)^2\right] \end{aligned}$$

Einsetzten von $f(x) = \phi(x) = x\log(x)$ gibt uns ein Resultat fuer die Entropie:

Korollar 4.43. *Sei $Y : \Omega \to [0, \infty)$ ZV, s.d. $E[\phi(Y)] < \infty$. Da ϕ konvex und diffbar ist:*

$f'(u) = (u \ln u)' = \ln u + \frac{u}{u}$

$$\mathrm{Ent}(Y) = \inf_{u>0} E\left[Y(\log Y - \log u) - (Y - u)\right]$$

$E(Y \ln Y - u \ln u - f'(u)(Y-u))$

$= E(Y \ln Y \; \cancel{- u \ln u} \; - Y \ln u \underbrace{- Y + \cancel{u \ln u} + u}_{-(Y-u)})$