

## Analys av signalsubstanser i hjärnan

### Bakgrund och målsättning

Projektet går ut på att studera frisättningen av dopamin hos nervceller. De två huvudsakliga frågeställningarna är hur osmotiskt tryck och kolesterohalt påverkar processen. Syftet med projektuppgiften är bland annat att ni ska träna på att

- använda Matlab för att tillämpa statistiska metoder på ett kemiskt forskningsproblem.
- konstruera och analysera statistiska modeller samt utföra kritisk granskning av modellernas förmåga att beskriva verkligheten.
- skriftligt redovisa modeller, antaganden och slutsatser från statistiska analyser.

Arbetet kommer genomföras i grupper om två till tre studenter. Projektet är indelat i fyra delar som motsvarar varsin datorövning. Arbetet kommer huvudsakligen utföras under datorövningarna men troligen kommer mer tid behövas för att slutföra och rapportera arbetet.

Projektet ska redovisas i form av en skriftlig rapport som lämnas in senast **2015-10-14**. Lämna in rapporten som en PDF och inkludera också koden som ni använde för att utföra beräkningarna. Organisera koden så att en fil `projekt.m` utför all analys. Skicka filerna till `david.bolin@chalmers.se`. Återlämning och eventuell korrigerings/komplettering görs vid datorövningen under den sista läsveckan.

I Appendix finns mer tips och instruktioner kring rapporteringen av projektet.

### Inledning

Vi har cirka 130 miljarder nervceller i hjärnan. Liksom andra typer av celler i kroppen består nervceller bland annat av en cellkärna med omgivande näringsvätska. Det som är unikt för nervcellen är att den kommunicerar med andra celler genom nervimpulser. Dessa impulser är en form av elektriska urladdningar som sprids i cellen och sedan går vidare genom utskott från cellkroppen. Nervcellen tar emot signaler via utskott som kallas dendriter och leder nervimpulserna till andra celler via ett utskott som kallas axon. Informationsöverföringen från en nervcell till en annan sker vid änden på axonen. Kontaktområdet där en axon möter en annan nervcell kallas för synaps och delas in i tre delar: En presynaptisk del i axonens ände, det synaptiska gapet och en postsynaptisk del i den andra nervcellen. Se Figur 1 för en illustration av en nervcell och en synaps.

När en nervimpuls når synapsen triggas synaptiska vesiklar fyllda med signalsubstans att frigöra sitt innehåll i det synaptiska gapet. Denna signalsubstans tas inom en tusendels sekund emot av särskilda mottagarmolekyler, så kallade receptorer, i den postsynaptiska delen. Om den mottagande cellen stimuleras med en tillräcklig mängd av signalämne uppstår en ny impuls, som förs vidare till nästa nervcell. Det finns flera typer av signalsubstanser som har olika funktioner. Några vanligt förekommande är glutamat, noradrenalin och dopamin.

Processen när signalsubstans utsöndras från nervcellens vesiklar vid cellmembranet kallas för exocytos. Även om mycket är känt kring denna process pågår bland annat forskning kring hur mycket av sitt innehåll vesiklarna frisläpper vid fusion med cellmembranet. Av intresse är också att studera hur olika faktorer påverkar frisättningen av signalsubstans, så som kolesterohalt och osmotiskt tryck.

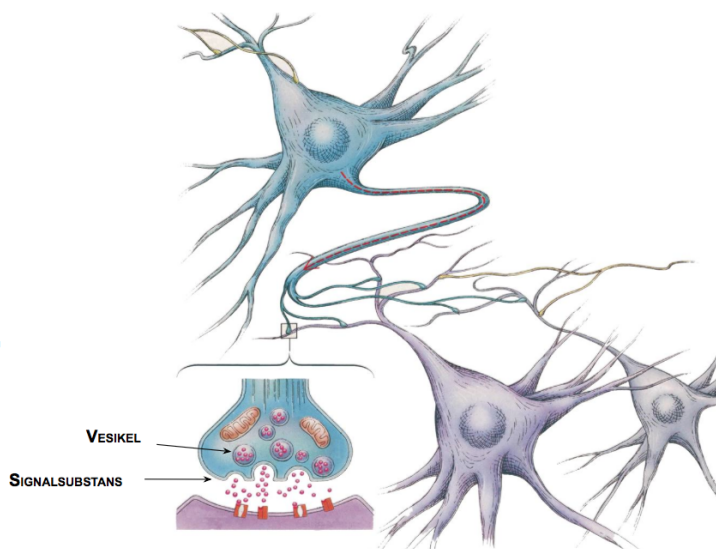


Figure 1: Illustration av nervcell och synaps. Källa: Fischback, G. D. Sci Amer. 1992, 267, 48-57.

## Amperometriska mätningar

För att studera hur frisättningen av dopamin går till kan amperometriska mätningar utföras. Mätningarna utförs genom att en microelektrod placeras mot membranet av en cell med vesiklar som kan frisläppa signalsubstans. När en vesikel frisläpper sitt innehåll registreras detta som en spik i signalen från elektroden, som illustreras i Figur 2. Signalen från elektroden analyseras och ett antal viktiga värden tas fram för varje spik i signalen.

Det första värdet är  $Q$ , vilket är arean under kurvan för spiken. Via Faradays lag kan vi relatera  $Q$  till antalet frisläppta molekyler,  $N$ :  $Q = nNF$ , där  $F$  är Faradays konstant och  $n$  är antalet elektroner som överförs per oxiderad molekyl (2 för dopamin).

Det andra viktiga måttet är  $t_{1/2}$ , vilket är tiden det tar från att spiken når halva sin höjd till att den faller ner under halva sin höjd igen. Detta mått, som brukar kallas halvtidsmätningar, säger alltså något om bredden på spiken och hur fort signalsubstansen frisläpps.

## Försöksuppställningar

Mätningar av  $Q$  och  $t_{1/2}$  har gjorts i två olika försöksuppställningar som kommer studeras i projektet.

Det första försöket går ut på att studera om osmotiskt tryck har någon påverkan på frisläppningen av dopamin. I försöket mäts först variablerna hos ett referensprov med normalt osmotiskt tryck och sedan mäts variablerna i ett prov med högt osmotiskt tryck. Dessa observationer kommer studeras i de första tre delarna av projektet.

Det andra försöket, som kommer studeras i den sista delen av projektet, går ut på att studera hur cellmembranets kolesterolhalt påverkar processen. Mätningar av processen har gjorts vid fyra olika nivåer av kolesterolhalt och dessa kommer studeras med hjälp av regressionsanalys.

## Datamaterialet

Varje grupp får en bit av det fullständiga datamaterialet att analysera. Datan ni ska analysera finns tillgänglig i filen `datagrxx` där `xx` är ert gruppnummer (grupp 4 hämtar alltså sin data från `datagr04`). Filerna finns tillgängliga på kurshemsidan. Om er grupp inte fått ett nummer än, skicka ett mail till [david.bolin@chalmers.se](mailto:david.bolin@chalmers.se) med namn och personnummer på deltagarna i gruppen.

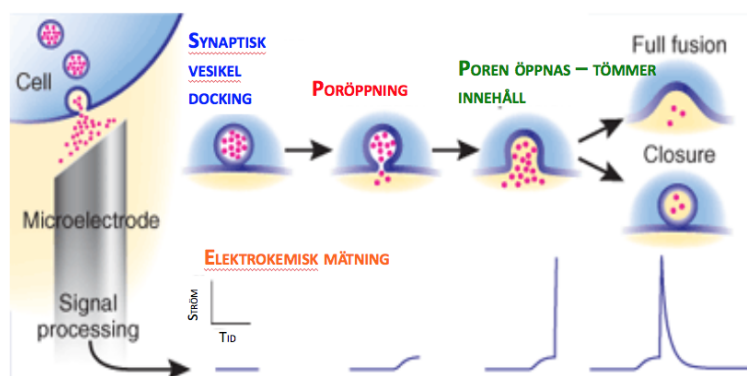


Figure 2: Illustration av processen då en vesikel frisätter signalsubstans samt den amperometriska mätningen av processen.

I datafilen finns resultaten från mätningarna relaterade till osmotiskt tryck sparade i variablerna

**t12nt** Uppmätta  $t_{1/2}$ -tider hos referensprovet med normalt osmotiskt tryck.

**Qnt** Uppmätta  $Q$ -värden hos provet med normalt osmotiskt tryck.

**t12ht** Uppmätta  $t_{1/2}$ -tider hos provet med högt osmotiskt tryck.

**Qht** Uppmätta  $Q$ -värden hos provet med högt osmotiskt tryck.

Resultaten från försöket med kolesterol finns sparade i variablerna

**t12cholxx** Uppmätta  $t_{1/2}$ -tider i prov med  $xx\%$  kolesterol.

**Qcholxx** Uppmätta  $Q$ -värden i prov med  $xx\%$  kolesterol.

De uppmätta värden på kolesterolhalten är  $xx = 0$ ,  $xx = 10$ ,  $xx = 20$  och  $xx = 30$ .

## Del 1: Undersökning av datan

Som alltid vid statistisk analys är det bra att börja med att göra en översikt av datan för att se om det finns anledning till oro kring hur mätningarna utförts. När vi senare gör en fördjupad analys kommer vi till exempel anta att mätningarna är oberoende och likafördelade, vilket bör undersökas. Börja med att titta på  $t_{1/2}$ -tiderna och  $Q$ -värdena för referensprovet med normal osmotiskt tryck och utför en preliminär analys:

- Plotta datan, gör histogram och beräkna t.ex. medelvärde och varians.
- Tycks det finnas några trender eller andra konstigheter i datan eller verkar mätsituationen varit under kontroll under hela studien?

Vi vill nu ansätta en statistisk modell för  $t_{1/2}$ -tiderna och  $Q$ -värdena och många olika fördelningar är tänkbara.

- Antag att vi som modell skulle använda normalfördelningen. Är detta en rimlig modell?
- Studera histogram av de observerade värdena, kan du hitta någon standardfördelning som passar bra till datan?

## Del 2: Modellanpassning

Av erfarenhet vet man att lognormalfördelningen kan passa bra till mätningar av detta slag. Slumpvariabeln  $X$  är lognormalfördelad med parametrar  $\mu$  och  $\sigma$  om  $\log X \sim N(\mu, \sigma^2)$ . Det innebär att  $X$  har täthetsfunktion

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\log(x) - \mu)^2\right), \quad x > 0.$$

Väntevärde och varians ges av  $E(X) = e^{\mu + \sigma^2/2}$  och  $V(X) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$ .

- Beräkna maximumlikelihood-skattarna av  $\mu$  och  $\sigma$  givet ett stickprov  $x_1, \dots, x_n$  från lognormalfördelningen. Ta med härledningen i er rapport.

Undersök nu om lognormalfördelningen är en rimlig modell för  $t_{1/2}$ -tiderna och Q-värdena:

- Undersök hur lognormalfördelningen ser ut genom att rita ut dess täthetsfunktion. Välj parametrarna enligt ML-skattningarna och jämför med histogrammet för datan. Använd också `normplot` för att undersöka om lognormalfördelningen passar bra, hur ska detta göras?
- Verkar lognormalfördelningen vara en rimlig modell för datan?

## Del 3: Undersökning av påverkan av osmotiskt tryck

Vi är nu intresserade av att testa hur osmotiskt tryck påverkar frisläppningen av signalsubstans. På grund av den tidigare analysen antar vi att det logaritmerade antalet följer en normalfördelning för både referensmätningarna och mätningarna undre högre tryck. Låt  $\mu_r$  och  $\mu_t$  beteckna väntevärdena för referensprovet och provet taget under högre tryck och låt  $\sigma_r^2$  och  $\sigma_t^2$  beteckna motsvarande varianser.

Börja med att testa om det är rimligt att anta att variansen för antalet är samma för de två försöken. Det vill säga utför hypotestestet

$$\begin{aligned} H_0 : \sigma_r &= \sigma_t \\ H_1 : \sigma_r &\neq \sigma_t \end{aligned}$$

med hjälp av ett F-test:

- Beräkna ett konfidensintervall för kvoten  $\sigma_r^2/\sigma_t^2$ .
- Beräkna p-värdet för testet. Kan nollhypotesen förkastas?

Testa nu om det osmotiska trycket påverkar väntevärdet hos fördelningen. Utför alltså hypotestestet

$$\begin{aligned} H_0 : \mu_r &= \mu_t \\ H_1 : \mu_r &\neq \mu_t \end{aligned}$$

Om ni inte kunde förkasta antagandet om lika varianser, utför testet med en poolad skattning av variansen. Utför annars testet när ni antar olika varianser. Kom ihåg att redovisa modeller och antaganden i rapporten.

- Beräkna p-värdet för testet, Kan nollhypotesen förkastas?
- Vad är slutsatserna av studien? Kommentarer?

Notera att resultaten i denna del av projektet ska redovisas både för  $t_{1/2}$ -tiderna och för Q-värdena.

## Del 4: Undersökning av kolesterols påverkan

Vi lämnar nu undersökningen av osmotiskt tryck och fokuserar på cellmembranets kolesterolhalt. I datafilen finns mätningar av  $Q$ -värden och  $t_{1/2}$ -tider för fyra olika halter av kolesterol. Med hjälp av dessa mätningar, utför en regressionsanalys av hur kolesterol påverkar processen:

- Från analysen ovan vet vi att både  $t_{1/2}$ -tider och  $Q$ -värden passar bra att modelleras med log-normalfördelningen. Ansätt regressionsmodeller med  $\log(t_{1/2})$  och  $\log(Q)$  som responsvariabler och kolesterolhalt som kovariat och skatta parametrarna.
- Undersök om regressionsmodellen passar bra genom att utföra lämplig residualanalys, plotta även datan och den skattade linjen.
- Skatta den förväntade förändringarna i  $t_{1/2}$  och  $Q$  om kolesterolhalten ökar med 1%. Beräkna även konfidensintervall för variablerna. Har kolesterol en signifikant påverkan på någon av variablerna?

## Appendix - Riktlinjer för projektredovisning

I denna sektion följer anvisningar som är bra att tänka på när ni skriver rapporten för projektarbetet.

Ni ska med stöd av frågorna i projekthandledningen skriva en självständig rapport som täcker det väsentliga innehållet. Målgruppen för skriften i projektet ska vara en teknolog i samma årskurs som läst kursen men som inte är insatt i detaljerna kring projektet. Språket i rapporten ska vara anpassat för målgruppen och texten ska vara tillräckligt omfattande och tydlig så att en person i målgruppen utan större ansträngning skall kunna följa resonemang och motiveringar. Speciellt måste texten vara korrekturläst så att språk- och skrivfel är rättade.

Rapporten ska innehålla klara och tydliga formuleringar av frågeställningarna, modeller och antaganden. Texten ska vara välstrukturerad med tydliga avsnittsrubriker och ska kunna läsas utan tillgång till vare sig kod som använts för beräkningar eller denna projektbeskrivning. Figurer och tabeller, försedda med figurtexter och tydlig numrering, ska användas för att redovisa resultat när det är lämpligt. Rapporten skall innehålla titelsida, sammanfattning och referenslista. Källhänvisningar ska vara enligt gängse norm men hänvisningar till kursboken och föreläsningssanteckningar får göras i förenklad form.

### Bedömning

Rapporten kommer i huvudsak bedömas enligt riktlinjerna ovan samt följande kriterier:

- att lämpligt och tydligt statistiskt språk används;
- att lämpliga tabeller, figurer och sammanfattande mått används för att beskriva datan och resultaten;
- att de statistiska beräkningarna är rätt utförda och verkar rimliga;
- att korrekta slutsatser dras med kommentarer om eventuella brister i analysen.

Nedan finns en checklista för granskning av rapporter. För att undvika onödiga misstag, gå igenom och bocka av listan med frågor innan rapporten lämnas in och se till att alla krav är uppfyllda.

## Checklista för rapportering av projektet

	Ja	Nej
1. Har alla moment i uppgiften blivit slutförda?	<input type="checkbox"/>	<input type="checkbox"/>
2. Innehåller rapporten relevanta figurer och tabeller?	<input type="checkbox"/>	<input type="checkbox"/>
3. Introduceras och förklaras all notation som används?	<input type="checkbox"/>	<input type="checkbox"/>
4. Introduceras alla modeller som används?	<input type="checkbox"/>	<input type="checkbox"/>
5. Presenteras och diskuteras resultaten tillräckligt?	<input type="checkbox"/>	<input type="checkbox"/>
6. Är alla beräkningar kontrollräknade?	<input type="checkbox"/>	<input type="checkbox"/>
7. Har rimligheten bedömts hos resultaten och har eventuella orimligheter kontrollerats och kommenterats?	<input type="checkbox"/>	<input type="checkbox"/>
8. Har rapporten: <ul style="list-style-type: none"> <li>• Titel, författare och datum?</li> <li>• Inledning?</li> <li>• Resultat och slutsatser?</li> </ul>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
9. Har rapporten korrekturlästs? Är stavningsfel rättade?	<input type="checkbox"/>	<input type="checkbox"/>
10. Angående figurer och tabeller: <ul style="list-style-type: none"> <li>• Är de numrerade?</li> <li>• Har de lämpliga figurtexter?</li> <li>• Är de refererade i texten?</li> </ul>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
11. Är texten välstrukturerad och indelad i stycken med lämpliga rubriker?	<input type="checkbox"/>	<input type="checkbox"/>
12. Går det att läsa och förstå rapporten utan: <ul style="list-style-type: none"> <li>• tillgång till projektbeskrivningen?</li> <li>• tillgång till Matlabkod som använts?</li> </ul>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
13. Om kursboken eller föreläsningarna refereras i texten, är referenserna tillräckligt specifika (anges kapitel/nummer på föreläsning)?	<input type="checkbox"/>	<input type="checkbox"/>