A bonyolult sztochasztikus rendszerek dinamikájának és strukturájának vizsgálata című pályázatunkban vizsgált problémák és bizonyított eredmények ismertetése.

Egyik fontos kutatási témánk a Wiener folyamatok és a lokális idő tanulmányozása volt. A 3 és magasabb dimenziós Wiener folyamat trajektóriája körüli r sugarú tartomány, az ún. "Wiener sausage" tulajdonságait vizsgáljuk. Ismert volt, hogy a tartomány térfogata eloszlásban közel (egydimenziós) Wiener folyamat, midőn az idő végtelenhez tart. A 3 dimenziós esetben erős approximációt adunk erre a közelítésre. (Csáki, E. and Hu, Y.: Strong approximations of three-dimensional Wiener sausages.)

Erdős és Taylor klasszikus eredményei a 3, ill. magasabb dimenziós bolyongás lokális idejének maximumára adnak 1 valószínűségű határértékeket. Két pont lokális idejének, ill. 1 pont lokális idejének és az e körüli gömb tartózkodási idejének együttes viselkedésére adunk 1 valószínűségű határérték tételeket. (Csáki, E., Földes, A., Révész, P.: Joint asymptotic behavior of local and occupation times of random walk in higher dimension.)

A Pólya tétel szerint 3, ill. magasabb dimenzióban a szimmetrikus bolyongás tranziens, azaz minden pontot csak véges sokszor látogat meg 1 valószínűséggel. Erdős és Taylor egy eredménye szerint azonban a látogatások számának maximuma n ideig $\log n$ nagyságrendű. Vizsgáltuk azon pontokat, ahol a bolyongás tartózkodási ideje a lehető legnagyobb, és bebizonyítottuk, hogy azok környezetében is az egyes pontok látogatásának száma aszimptotikusan konstansszor $\log n$, midőn n-nel végtelenhez tartunk. (Csáki, E., Földes, A. and Révész, P.: On the behavior of random walk around heavy points.)

Egy másik a fenti problémakörrel rokon fontos kutatási témánk a véletlen bolyongások egy fontos, de nehezen vizsgálható mennyiségének, a lokális időnek a viselkedése volt.

Vizsgáltuk az aszimmetrikus bolyongás lokális idejét abban az esetben, amikor a számegyenesen egy pontból a szomszédos pontokba lépés valószínűségei p, ill. $q = 1 - p$, $p > q$. Ismeretes, hogy ez a bolyongás tranziens, a lokális idő a magasabb dimenziós bolyongáséhoz hasonló tulajdonságokat mutat, azaz n lépésben a maximális lokális idő $\log n$ nagyságrendű, és az ilyen maximális lokális idejű pontok környezetében a lokális idő nagyságrendje ugyancsak $\log n$. Ezt a jelenséget pontosítottuk, majdnem biztos határértékeket határoztunk meg pontos konstansokkal, illetve azt vizsgáltuk, hogy milyen nagy környezetben igaz még ez a jelenség. (Csáki, E., Földes, A. and Révész, P.: On the local time of the asymmetric Bernoulli walk.)

Ezenkívül olyan bolyongásokat is vizsgáltunk a számegyenesen, amelyekben az egy pontból a szomszéd pontokba lépés valószínűsége függ a helytől, de határértékben ezek a valószínűségek 1/2-hez tartanak. A valószínűségektől függően egy ilyen bolyongás bizonyos esetekben rekurrens, más esetekben tranziens. A tranziens esetet vizsgáltuk, és meghatároztuk a lokális idő eloszlását, illetve annak aszimptotikáját, amikor a hely "kimegy a végtelenbe". Meghatároztuk az olyan szakaszok hosszának az aszimptotikáját, amelyek során a bolyongás csak előre lép. Csáki, E., Földes, A. and Révész, P.: Transient nearest neighbor random walk on the line.)

További eredményeket is elértünk az ilyen modellek vizsgálatában. A bolyongás és a megfelelő Bessel folyamat között erős approximációt bizonyítottunk. Vizsgáltuk, hogy a Bessel folyamat majdnem biztos tulajdonságai ilyen módon mennyiben öröklődnek át a bolyongásra. Meghatároztuk, hogy az átmenet valószínűségektől függően milyen esetben van véges, illetve végtelen sok elválasztó pont (cutpoint), azaz olyan pont, amely a bolyongás első felét elválasztja a másodiktól. (Ez azt jelenti, hogy a két rész diszjunkt). (Csáki, E., Földes, A. and Révész, P.: On the number of cutpoints of the transient nearest neighbor random walk on the line.)

Közönséges bolyongás lokális idejének növekményeire erős approximációt adtunk egy kétparaméteres Gauss folyamat segítségével. Ennek következményeként határeloszlástételeket, illetve majdnem biztos határérték tételeket tudtunk bizonyítani a lokális idő növekményeire. (Csáki, E., Csörgő, M., Földes, A. and Révész, P.: Random walk local time approximated by a Brownian sheet combined with an independent Brownian motion.)

Foglalkoztunk független valószínűségi változók viselkedését leíró eredmények természetes általánosításaival. Egyik dogozatunk a független valószínűségi változók normált részletöszegei maximumára vonatkozó Darling-Erdős tétel pontonkénti változatát írja le, és egy olyan paraméteres határeloszlástétel-osztályt ad meg, ahol a pontonkénti változatban fellépő átlagolási eljárás folytonosan változik a log és $\log\log$ átlagolás között. (Berkes I. Weber, M.: Almost sure versions of the Darling-Erdős theorem.) Független, azonos eloszlású valószínűségi változók súlyozott részletösszegeire bizonyítottunk iterált logaritmus tételeket. Ha a változókra csak 2 momentumot teszünk fel, az aszimptotika nem az együtthatók nagyságrendjétől, hanem az együtthatósorozat aritmetikai tulajdonságaitól függ. Tanulmányoztuk a problémát néhány klasszikus "szabálytalan" együtthatósorozat (pl. additív számelméleti függvények) esetében. (Berkes I., Philipp, W., Tichy, R.: Empirical processes in probabilistic number theory: the LIL for the discrepancy of $(n_k\omega)$ mod 1.) Továbbá független valószínűségi változók összegeinek un. "teljes" konvergenciájára adtunk új feltételeket.

Más dolgozatokban a számelméletben fellépő sztochasztikus jelenségekkel foglalkoznak. Az additív számelméleti függvények eloszlására vonatkozó Erdős-Kac féle centrális határeloszlástételnek megfelelő iterált logaritmus tételt bizonyítottuk be. (Berkes I., Weber M.: Moment convergence and the law of the iterated logarithm for additive functions.)

E kérdés megoldását a pontonkénti CHT elmélet jelenségeinek vizsgálata tette lehetővé. Megadtuk az $\{n_k\alpha\}$ típusú sorozatok diszkrepanciájának pontos aszimptotikáját hézagos, de exponenciálisnál lassabban növő n_k sorozatok esetén. Ez az aszimptotika erősen függ az n_k sorozat számelméleti (diophantoszi) tulajdonságaitól. Megadtuk nemlineáris sorozatok diszkrepanciájának pontos aszimptotikáját hézagos, de exponenciálisnál lassabban növő n_k sorozat esetén. Ez az aszimptotika erősen függ az n_k sorozat számelméleti (diophantoszi) tulajdonságaitól. Foglalkoztunk nemlineáris idősorok aszimptotikus és statisztikai tulajdonságaival, nevezetesen az ilyen folyamatok paraméterváltozásainak észlelésére, valamint bizonyos nemlineáris funkcionáljai és lokális ideje aszimptotikájára vonatkozó eredményeket bizonyítottunk.

Vizsgáltuk azt a kérdést, hogy a klasszikus valószínűségszámítás eredményei független valószínűségi változók összegeire hogyan változnak meg, ha a függetlenség helyett valamilyen más, csak gyenge függőséget biztosító tulajdonság érvényesül.

Így például Berkes I. és Aistleitner C. On the central limit theorem for $f(n_k x)$ című dolgozata szükséges és elégséges feltételt fogalmaz meg arra, hogy egész számok egy (n_k) sorozata és egy sima periodikus f függvény esetén a $\sum_{k=1}^{N} f(n_k x)$ összegekre igaz legyen a centrális határeloszlástétel. E problémát M. Kac vetette fel 1949-ben, és a mai napig megoldatlan maradt. Kac megmutatta, hogy az $n_k = 2^k$ esetben érvényes a CHT, és nem sokkal később Erdős and Fortet igazolták, hogy az $n_k = 2^k - 1$ esetben ez nem igaz. Ez azt sugallja, hogy a CHT érvényessége az (n_k) számelméleti tulajdonságain múlik. Az említett dolgozatban megmutattuk, hogy a CHT akkor és csak akkor teljesül, ha az $an_k + bn_l = c$, $1 \leq k, l \leq N$ diophantoszi egyenlet megoldásainak száma "nem túl" nagy.

Berkes I., Hörman S. Horváth L. az Asymptotic results for the empirical process of stationary sequences című dolgozatban egy új keverési együtthatót ("S-mixing") vezettek be, amelynek érvényessége a szokásos keverési együtthatóknál (erős keverés, ϕ-keverés, stb.) jóval egyszerűbben ellenőrizhető, és aszimptotikus következményei ugyanolyan erősek, mint a klasszikus keverési fogalmaké. Ennek felhasználásával stacionárius folyamatok empirikus eloszlásfüggvényére bizonyítottak be új eredményeket. Egy másik dolgozatban az arcus sinus törvény funkcionális formáival foglalkoztunk.

Más dolgozatokban néhány fontos statisztikai problémát vizsgáltunk. Így például foglalkoztunk az un. "change point" kérdéskörrel. Itt egy folyamat struktúrájában fellépő változásokra próbálunk következtetni egy minta alapján. Egyrészt lineáris (azaz mozgó átlag) folyamatok kovariancia-struktúrájának változására konstruáltunk egy tesztet. (Berkes I., Gombay E, Horváth L. Testing for the changes in the covariance structure of linear processes.) Egy másik munkánkban az un. funkcionális adathalmazokban fellépő strukturális változásokkal foglalkoztunk. (Berkes I., Gabrys R., Horváth L., Kokoszka P.: Detecting for changes in the mean of functional observations.) Funkcionális adathalmazok többek között az ekonometriában lépnek fel, amikor "nagy sűrűségű" megfigyelésekkel dolgozunk, pl. egy részvény árát egy több hónapos időintervallumban percenként ismerjük. Ilyenkor az n-edik adat a részvény n-edik napon megfigyelt értékeiből áll, amelyeket nem egy magas dimenziójú vektornak tekintünk, hanem egy függvénynek, és az adatok analízise absztrakt tér (pl. Hilbert tér) értékű valószínűségi változók vizsgálatát kívánja meg.

Más vizsgálatokban véges halmazokból való mintavételre bizonyítottunk új típusú eredményeket. Ha a halmazok elemei teljesítik a szokásos "aszimptotikus elhanyagolhatósági" feltétetelt, akkor a mintaelemek összegére érvényes a centrális határeloszlástétel, akár visszatevéssel, akár visszatevés nélkül húzunk. (Berkes I., Hörmann S., L. Horváth: The functional central limit theorem for a family of GARCH observations with applications.) Egy másik dolgozatban leírtuk a részletösszegek viselkedését abban az esetben, ha az aszimptotikus elhanyagolhatósági feltétel nem teljesül. (Berkes I., Aue A., Horváth L.: Selection from a stable box.) Ilyenkor a mintavételre használt halmaz elemei valószínűségi változók eredményeink kiadják a végtelen szórású változókból vett minták bootstrap illetve permutáció statisztikáinak aszimptotikus viselkedését, amely a klasszikus viselkedéstől lényegesen eltér. A határeloszlás nem normális, és véletlen paramétereket tartalmaz.

Egy más kutatási témában független valószínűségi változók nem lineáris funkcionáljaival foglalkoztunk. Az U-statisztikák eloszlására adtunk olyan éles becsléseket, amelyek későbbi statisztikai vizsgálatokban is hasznosnak bizonyulhatnak. (Major P.: On a multivariate version of Bernstein's inequality.) Ezenkívül egy másik klasszikus egyenlőtlenségnek is megadtuk egy nem-triviális többváltozós általánosítását. (Major P.: On the tail behaviour of multiple random integrals and degenerate U-statistics.)

Foglalkoztunk a matematikai statisztika orvosi és biológiai alkalmazásaival is. Az ilyen jellegű kérdések néhány a valószínűségszámításban és matematikai statisztikában népszerű hálozatokkal kapcsolatos probléma vizsgálatához is elvezettek.

Véletlen struktúrák matematikai statisztikai vizsgálatában szerintünk az elsődleges szempont azt a funkciót modellezni, amelynek ellátására a struktúra létrejött. Véletlen hálózatok szerkezete azt a folyamatot irányítja, amely a csúcsok együttes állapotából áll. Ezen az alapon egyesítjük Erdős Pál és Rényi Alfréd gráfok evolúciójára, Albert Réka és Barabási Albert-László hálózatok preferencia alapú fejlődésére és Stuart Alan Kauffman enzimatikus kölcsönhatásokra vonatkozó modelljeit.

Bizonyos statisztikai alkalmazások felvetettek a véletlen permutációk finomabb tulajdonságairól szóló kérdéseket. Ilyen jellegű kutatásokat is folytattunk. Kerestük véletlen permutációk olyan feltételes függetlenségi relációkkal definiálható modelljeit, amelyek a klasszikus hierarchikus, illetve log-lineáris modellek analógjai a véletlen permutációk világában. Több (egyszerűbb és bonyolultabb) modellt sikerült megalkotni, amelyek a gyakorlatban is alkalmazhatóak, ugyanakkor szép matematikai tulajdonságokkal rendelkeznek. A hagyományos módszereken túl az algebrai statisztika eszköztára is jól alkalmazható ezen modellek vizsgálatánál. Igazoltuk P. McCullagh egy sejtését, amely egy általa definiált, véletlen permutációkra adott modell identifikálhatóságáról szólt. (Csiszár V. Markov bases for conditional independence models for permutations.)

Egy másik általunk vizsgált véletlen permutációkkal kapcsolatos probléma Persi Diaconisnak az Amerikai Pszichiáterek elnökségi választás alapján összeállított nagy elemszámú mintájához kapcsolódik. Ez egy az irodalomban jól ismert minta, amely azóta a véletlen permutációk statisztikai vizsgálatában etalonná vált. Sok modellt illesztettek az adatokra, de egyik illeszkedése sem volt elfogható. Az adatok elemzése vezetett egy egészen új modell felfedezésére. Ebben a modellben a legnehezebb feladat a szabadsági fok meghatározása volt, amelyhez a statisztikai modellek elméletének jelentős fejlesztésére volt szükség, amely elvezetett az algebrai statisztikához. Erről bizonyított érdekes eredményeket Csiszár Villő a Conditional independence relations and log-linear models for random matchings dolgozatában.

Évek óta tanulmányozunk egy 300 beteget tartalmazó 30 éve tartó pszichiátriai adatmezőt, egy ideje az eredményeket ismertető cikkeken dolgozunk. Kimutattuk, hogy a lelki zavarokkal küzdő betegek körébben az intelligencia enyhén alacsonyabb az átlagnál. A munka során egy-egy beteget több orvos is megvizsgált, a diagnózisok összehasonlítására Markov-modellt dolgoztunk ki, melyben az orvosoknak a betegekhez hasonlóan saját arculatuk van és a két állapottér kölcsönhatása az orvos véleménye a betegről.

Az intezétünkben működő csoport folyamatosan fejleszt egy univerzális modellt amely az immunológiai rendszer folyamatait írja le. A modell alapfeltevése, hogy az immun rendszer elsődleges feladata az egészséges szervezet harmóniájának a fenntartása, az idegen anyag kilökése ennek korolláriuma.

Foglalkoztunk további a biostatisztika és matematikai statisztika által felvetett problémákkal. Ilyen volt az enzím hálózatok vizsgálata.

Start Alan Kauffman kölcsönható Boole-hálózatok modelljét fejlesztették úgy hogy az enzimek aktivitását mérő úgy nevezett micro-array kisérlet eredményeit használták fel. A kísérleti adatokhoz való jó kvantitatív illeszkededésen túlmenően a Kauffman által meghatározott kritikus dinamikus viselkedést tették a modell elfogadhatóságának kvalitatív feltételévé. A kísérleti eredmények gondos mérlegelése a Kauffman és munkatársai által javasolt kanalizáló Boole függvények általánosítására vezetett. Azok a Boole függvények lettek a legjobban alkalmazhatóak, amelyek fa-reprezentációjában minden egyes változó csak egyszer fordul elő. (Tusnády G., Rejtő L. Reconstruction of Kauffman networks applying trees.)

Továbbá foglalkoztunk immunológiai és populációgenetikai modellekkel.

A matematikai immunológia hosszú idők óta megoldhatatlan kérdése az, hogyan dönti el az immun rendszer, mikor mi a teendő. Rengeteg matematikai modell született, de ezek keletkezése nem áll meg annak megfelelően, hogy mindegyik csak erős elhanyagolással kezeli a kérdést és az újabb eredmények érvénytelenítik. A Rényi Intézetben működő csoport egyre közelebb jut egy olyan teljes elmélet megteremtéséhez melyben minden jelenség értelmezése fehérje szinten történik. Az alapfeltevés az, hogy az immun rendszer fő feladata az egészséges szervezett életének koordinálása. (Tusnády G., Bakács T., Mehrishi J.N., Szabados T., Varga L., Szabó M.: T-cells survey the stability of the self: a testable hypothesis on th homeostatic role of TCR–MHC).

Az alábbi kérdés vezetetet populációgeneteikai modellek vizsgálatához.

Diszkrét időben fejlődő genetikai modellek között már 1995-ben találtak olyan példát mely konvergencia helyett kaotikus rendszerre vezet. Most sikerült először a jelenséget folytonos idejű folyamatokban rekonstruálni. (Tusnády G., Hatvani L. Tokos G.: A mutation-selection recombination model in population genetics.)

Végül foglalkoztunk gyakorlati statisztikai problámákkal is. Ilyen az alább ismertetett munka, ahol pszichiátriai adatok feldolgozásával foglalkoztunk.

Pethő Bertalan 1965 és 1975 között tíz nozológiai csoporthoz tartozó betegekből egy 250 fős mintát vizsgált meg tíz különböző teszttel amelyekben általában húsz pontozható tünet szerepelt. Körülbelül 50 fős normál kontroll csoporton is elvégezte ezeket a vizsgálatot. Ezt a mintát 1995 és 2000 között újra megvizsgálta. Az egyesített adatokat a Rényi Intézet dolgozta fel. A feldolgozáshoz a Boole hálózatokhoz hasonló új modellt dolgoztak ki amely feltételezi, hogy a vizsgált személyekben is és a vizsgálókban is rejtett faktorok vannak és a teszt eredménye a két komponen kölcsönhatása. A statisztikai eljárás az úgy nevezett EM algoritmus alkalmazása. A legfontosabb eredmények:

A) A nozológia csoprtok súllyosságukat tekintve a szakemberek által feltételezett módon rendezhetőek, egyesekben teljes gyógyulás várható, mások viszont gyakorlatilag gyógyíthatlanok, csak az állapotuk stabikizálására lehet törekedni.

B) Előfordul, hogy a besorolás megváltozik: ez nem szakmai hiba, hanem az élet rendje. Egy homokórára emlékeztető modell írja le a jelenséget a legplasztikusabban: sokszor a tünetek homályos és kaotikus együttese koncentrálódik, ekkor manifesztálódik a betegség, aztán az idő haladtával a tünek koncentráltsága, egyetlen nozológiai besorolásra utalása újra elmosódik.

C) A betegek csoportosítására két rendszer van használatban. A minta elemeit mind a két rendszerbe beillesztették, emiatt sikerült a kát rendszer kapcsolatát felderíteni.

D) Meghatározták az egyes tesztekre a vizsgálók döntésinek a megbizhatóságát (reliability).

E) Bebizonyították, hogy az egészségesek és betegek közötti határvonal egyértelműen meghúzható. Tusnády G., Pethő B., Vargha A., Tolna M., Farkas Gy., Vizkelety A., Tóth A., Szilágyi I., Bitter I., Kelemen A., Czobor P.: Validity and reliability; comparision of inter-rater reliabilities of psychopathological symptoms.)